

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images,

please do not report the images to the
Image Problem Mailbox.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2002, 00:48:08 ; Search time 9448.45 Seconds
(without alignments)
17973.196 Million cell updates/sec

Title: US-09-778-516a-1
Perfect score: 8115
Sequence: 1 gatgtacggcagatatac.....gcgtgtcgtacgacgctc 8115

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank:
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_cm:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgt_hum:*
31: em_hgt_inv:*
32: em_hgt_other:*
33: em_hgt_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
------------	-------	-------	-------	--------	----	-------------

1	3535	43.6	3660	1	LBAGAL	M23530 L. delbrueckii
2	3497.2	43.1	5059	6	AX009488	AX009488 Sequence
3	3386.6	41.7	5015	1	LBACACZBUL	M55068 L. delbrueckii
4	1336.4	16.5	3840	12	EVEI32038	AJ132038 Expressio
5	1307.4	16.1	11592	6	AX107924	AX107924 Sequence
6	1245.2	15.3	3853	6	BD009729	BD009729 Tissue sp
7	1240.4	15.2	3853	6	AR098190	AR098190 Sequence
8	1236.2	15.2	2093	1	LP1PLPRP	M1223 Plasmid PLP
9	1191.6	14.7	2140	1	C300RFR	J03319 Plasmid PC3
10	1152.4	14.2	4699	12	AF346633	AF346633 RAGE vect
11	1148.2	14.1	14455	12	AF356044	AF356044 Sequence
12	1082	13.3	3986	12	PCDMA32EO	X90639 Cloning vec
13	1082	13.3	5446	6	AX319694	AX319694 Sequence
14	1068.6	13.2	4026	6	AR098191	AR098191 Sequence
15	1067	13.1	4026	6	BD009730	BD009730 Tissue sp
16	1063.4	13.1	2821	6	AX195202	AX195202 Sequence
17	1063.4	13.1	2833	6	AX195203	AX195203 Sequence
18	1053.4	13.1	2845	6	AX195204	AX195204 Sequence
19	1058.4	13.0	4622	6	AX299128	AX299128 Sequence
20	1036.2	12.8	3796	6	E49323	E49323 Infectious
21	990	12.2	6299	6	AX164584	AX164584 Sequence
22	988	12.2	4597	6	AX060344	AX060344 Sequence
23	985.2	12.1	4276	6	AX154496	AX154496 Sequence
24	985.2	12.1	4282	6	AX164583	AX164583 Sequence
25	985.2	12.1	5128	6	AX154497	AX154497 Sequence
26	985.2	12.1	5459	6	AX154499	AX154499 Sequence
27	985.2	12.1	5882	6	AX154501	AX154501 Sequence
28	985.2	12.1	9620	6	AX164579	AX164579 Sequence
29	965.2	12.1	9620	6	AX164581	AX164581 Sequence
30	973	12.0	4283	6	163120	163120 Sequence 3
31	973	12.0	4283	6	185496	185496 Sequence 1
32	968.6	11.9	4050	12	AF053407	AF053407 Expressio
33	968.6	11.9	4227	12	AF053408	AF053408 Expressio
34	968.6	11.9	4625	12	AF053409	AF053409 Expressio
35	968.2	11.9	4657	12	AF053406	AF053406 Expressio
36	961.4	11.8	6200	6	AX353687	AX353687 Sequence
37	959	11.8	5732	6	AX027904	AX027904 Sequence
38	959	11.8	5732	12	AX027910	AX027910 Sequence
39	935.6	11.5	5432	6	AX026821	AX026821 Sequence
40	935.2	11.5	5729	6	AX027908	AX027908 Sequence
41	935.2	11.5	5729	12	AX027914	AX027914 Sequence
42	933.4	11.5	4392	12	AF361441	AF361441 Cloning v
43	928.6	11.4	6094	6	AX027906	AX027906 Sequence
44	928.6	11.4	6094	12	AX027912	AX027912 Sequence
45	921.8	11.4	6285	6	AR018923	AR018923 Sequence

ALIGNMENTS

RESULT 1	LBAGAL	3660 bp	DNA	linear	BCT 26-APR-1993
LOCUS	LBAGAL				
DEFINITION	L. delbrueckii bulgaricus beta-galactosidase gene, complete cds.				
ACCESSION	M23530				
VERSION	M23530.1 GI:149546				
KEYWORDS	galactosidase.				
SOURCE	L. delbrueckii bulgaricus (strain BJ31) DNA.				
ORGANISM	Lactobacillus delbrueckii				
REFERENCE	1 (bases 1 to 3660)				
AUTHORS	Schmidt, B.F., Adams, R.M., Regnadt, C., Power, S. and Mainzer, S.E.				
TITLE	Expression and nucleotide sequence of the Lactobacillus bulgaricus				
JOURNAL	beta-galactosidase gene cloned in Escherichia coli				
MEDLINE	J. Bacteriol. 171, 625-635 (1989)				
COMMENT	Draft entry and computer-readable sequence for [1] kindly submitted				
FEATURES	by B.F. Schmidt, 29-MAR-1989.				
Source	Location/Qualifiers				
	1..3660				
	/organism="Lactobacillus delbrueckii"				
	/db_xref="taxon:1584"				

Db 1586 CATGACGAGCAACATCAATGCTGTCCGCTGCTCTCACTACCGCAACGAGTCCCTTT 1645
Oy 6160 ttaccggctctgtgacaagaftacggcctttactgcatgtaagactaacctgaaagcca 6219
Db 1646 TTACCGGCTCTGTGACAAAGTACGGCTTTACGTCATTGATGAAGCTAACCTGGAAACCA 1705
Oy 6220 cggcaccctgggaaaaagctgggggggacagaaagccttaagcttcaatgttccaagcgtga 6279
Db 1706 CGGACCTGGGAAAAAAGTGGGGGGGCGACGAAGATCCAGCTTCATGTTTCAGGCGATGA 1765
Oy 6280 ccagcatctgctgggacagcttaccgggtgaaagaaatgaatgctcgggaaagaa 6339
Db 1766 CCAGCATGGCTGGAGCCAGCTTATCCGGGTGAAGAAATGATGCTCGGACAGAA 1825
Oy 6340 ccatgtcctaactaaatcgtgctttagcaatgagcttaagccggcactgtcttgc 6399
Db 1826 CCAATGGCTGATTTAGCTGGCTTTAGGCAATGATGCTTACGCCGCGACTGCTTTC 1885
Oy 6400 ccaatgtgctatcagctccggaagctgaaccgaccgggttcaagcactaagaaggtc 6459
Db 1886 CCAATGGCTGATTTAGCTGGGAGGCTGATCCGACCCGGGTTCCAGACTATGAAGGCT 1945
Oy 6460 gaccacaacaggaagcttgaagcagccacagattgaagccggaatgatagtccgctc 6519
Db 1946 GACCCCAACCGGAAGTTTGAAGCAGCCAGATGAAGCCGAGTATGCTCCGCG 2005
Oy 6520 caaggaattggaagtaacttgacccaataacagcccaagccattatctcagttgaata 6579
Db 2006 CAAGGTAATTTGAAGTAATCTTGACCAATAAACAGCCCAACCAATTTATCTCAGTTGAATA 2065
Oy 6580 cgtctacagcattggaacactcgtctggtgacttgcgacctacacggcccttggaaata 6639
Db 2066 CGCTCAAGCCATGGGCAACTCGTGGTGACCTGGCCGCTTACAGCGCCCTGGAAAAATA 2125
Oy 6640 ccccaactacagagcggtctcatctggagctgattgaccaaagcattgaaagaaagcgg 6699
Db 2126 CCCCCATTCACAGGGGGGCTTCATCTGGGACTGGATTGACCAAGAGATCGTGAAGAAAGACGG 2185
Oy 6700 gcaacgcttattgaggggagcacttgatgacggccaacagcactatgaattctcggggaa 6759
Db 2186 GCACCTGCTTTATGGGGGCGACTTCGATGACCGGCCCAACCGACTATGAATTCGCGGAA 2245
Oy 6760 cggcctgtctctgtctgacccggaactgaatcgcggaacttgcataatgcaagccctta 6819
Db 2246 CGGCTGTGCTTTGTGACCGGAGCTGATGCGGAAACCTGGTAATGTCAGGCCCTTTA 2305
Oy 6820 cggcaaccttaagtaagaagtaaaagatggagcctctctcctaagaaagcaattatct 6879
Db 2306 CGCCCACTTAATGTAAGATGAAGATGGGCACTCTTCTCAAAAACACCAATTTATT 2365
Oy 6880 taaccaacgctcactctactactctctgactagcttcttggctgcatggcaagttgacct 6939
Db 2366 TACCAACAGCTCATTTACTACTCTTGTGACTATCTTTTGTGTGATGGCAAGTTGACTTA 2425
Oy 6940 ccagagccggcctctgacacttggccttgagccttggcgaatccgggaacttgccttgc 6999
Db 2426 CCAGAGCCGGCTCTGACCTTTGGCTGGAGCTGGGGAATCCGGGACCTTCCCTGCC 2485
Oy 7000 ttggccgggaagtcgctgatagaagaaagggagtcgctctacgggttaacggccacttaa 7059
Db 2486 TTGGCCGGGAAGTGGTGAATGAAGAAAGGGAGGTCGTCTACCGGGTAAACGCCCACTTAA 2545
Oy 7060 agaagacttgacctggcggaatgagggtcactgtgactggaagcagaagaagtactca 7119
Db 2546 AGAAGACTTGCTTTGGGCGGATGAGGGCTTCACTGTGGCTGAAGCGAAGAAATAGCTCA 2605
Oy 7120 aaagtgcgggaatttaagccggaagggcgagcaagattagttgattccgactacaact 7179
Db 2606 AAAGCTGCGGGAATTTAAGCGGGAAGGGCGGACAGATTTAGTTGATTCGACTACAACT 2665
Oy 7180 aggcctgaaagaaataacttcaaatctctctcgaaggtcaagggctgcgggttc 7239
Db 2666 AGGCTGAAGGAATTAATCTTCAAAATCTCTTCCAAAGGTCAAGGGCTGGCGGTTTC 2725

Oy 7240 cctcaagtatgcccgttagggaaataactgaaagcggctccggaattactcttgcgggc 7299
Db 2726 CCTCAAGTATGCGGATGGGATATCTTGAAGCGGCTCCGGAAATTTACCTTCTGGCGGC 2785
Oy 7300 cctgacgacaacagcaccgggagctgttgaagctatgacttgcgcccgttgggaaatgc 7359
Db 2786 CTTAGCGACACAGCACCGGGGAGCTGTACGGCTATGATCTGCGCCGGTGGGAAATGC 2845
Oy 7360 cggcgaatgcccgttgaagaaacatcagctgcgaagtcgaagaaagactcgttgt 7419
Db 2846 CGGCAAGTATGCCCGCTTGAAGACATACACTGCGAGGTCAAGGAACATCCGTTTGGT 2905
Oy 7420 caagagcctttagcttgccttgccttgccttgaagggatgatttaacgctgactaagtc 7479
Db 2906 CAAGACTGCTTTAAGTGTGCTGCTGCGCTTAAAGGATGATTTAACTGCTGACTTAAGT 2965
Oy 7480 cgaatgacgggcaagattgtctgaacagctgaacttccagcgcggaagaagctgtct 7539
Db 2966 CGATGAGAGGGGCAAGATTCTGTAAACAGCTGACTTCCAGGGCGGGAAGAGCTGCT 3025
Oy 7540 ctgacagccttgccttgaacctgacctgcctgcgaagaaactgaacgactaacta 7599
Db 3026 CTGCGCAGCTTTGGCTTGAACCTGCGCTGCGCAAAAGAACTGACCGATTACCGCTACTA 3085
Oy 7600 tggcttggagacttaatgagagctaccagacgcttggaaagatgaattactctggacata 7659
Db 3086 TGGCTTGGACCTTAATGAGCTAACCCAGACCGCTTGGAAAGTAATTAATCTGGGCACTTA 3145
Oy 7660 ccagggagcgttaaaaaaagaacttgaacccatctgcgcgagaaagcggaacggag 7719
Db 3146 CCAGGAGGGGTTAAAAAAGAACTTACCATATGCTGCGCAGAAAGGGCAACCGGAG 3205
Oy 7720 caaggttcgtgttaccagctcttgaatgaagggcggttggaaattacggccaatg 7779
Db 3206 CAAGGTTCGCTGTGACAGCTCTTGTGAAGAAAGGGCGGCTTGAATTTACGCCAATG 3265
Oy 7780 ggcagacttgaacttgccttgccttgccttgccttgccttgccttgccttgccttgc 7839
Db 3266 GGCAGACTTAATCTTGTCTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 3325
Oy 7840 cgtcttgaacttgaacttgaacttgaacttgaacttgaacttgaacttgaacttgaact 7899
Db 3326 CGCTTTGAACCTGATTAACATTAATGATGCTTGAAGGCTTGAAGGCGCCAGATGGGG 3385
Oy 7900 cggcggggatgactccttgggggacagaagctccacccggaattctgcctgatagtctaaa 7959
Db 3386 CGGCGGGGATGACTCTGGGGGACAGAGTCCACCCGGAATTCGCTGATGCTCAAAA 3445
Oy 7960 agcccgcaagcttgccttgccttgccttgccttgccttgccttgccttgccttgcct 8019
Db 3446 AGCCCGGACCTTGCCTGTGATTCAGCCCTTTTACTTAATAATGCTACAAATTTGAC 3505
Oy 8020 ttaacagatgaatlttagttaaagcaaaagcagatgaggaagatgycacaagatcagaga 8079
Db 3506 TTAACAGATGAATTTTACTTAAGCAAAAGCAGATGAGGAGATGSCAAGATCAGAGA 3565
Oy 8080 agtgcgaagcagcgcgctgtcgtcagcagcgt 8114
Db 3566 AGTGCACAGGACACCGCGGTGCTAGGCAAGCGT 3600

RESULT 2
AX009488 5059 bp DNA linear PAT 06-SEP-2000
LOCUS AX009488
DEFINITION Sequence 1 from Patent WO961627.
ACCESSION AX009488
VERSION AX009488.1 GI:9996771
KEYWORDS
ORGANISM
Lactobacillus delbrueckii subsp. bulgaricus.
Lactobacillus delbrueckii subsp. bulgaricus
Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
Lactobacillus.

REFERENCE 1 (bases 1 to 5059)
ATTHONS Brignon, P., Gendreau, F., and Benbadis, L.
TITLE Mutant *Lactobacillus bulgaricus* strains free from beta-galactoside
JOURNAL activity
Patient: WO 9961627-A-1 02-DEC-1999;
GERVAIS DANONE CO (FR); BRIGNON PIERRE (FR); GENDRE FRANCOIS (FR);
BENBADIS LAURENT (FR)
FEATURES
source location/Ouallifiers
1. 5059
/organism="Lactobacillus delbrueckii subsp. bulgaricus"
/db_xref="taxon:1585"
122..1873
CDS
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07487.1"
/db_xref="GI:9996773"
/translation="MIFITNLITAIRIGEVLDPLIGNAIDRTESRWKFKPWWVGG
GIISLALALFTDFGGINOSNPVYLVIFGIYVLMIDIFESFKDTEFWMIPLSID
SREBKTSTFARVGSITIGANLVGVITPILIFESKRNANGDQKQGFALIVATIG
LITSTVIGLSTHEKRSALRESNEKTLOKVEKVGONDOLMLAFWYFGLITNG
ALDIYFSTILGDRGSLTYTITFVGLISAFSPSLAKFNENRLEFACIAWMLG
IGVSVASGLASINGAEFFIIPDLAPLIVMLITSDAYEGOLKTHGDEALTLISV
RPLVDKLGALSNNFVSLILTAOMTTGATASTITAGOMVFKLAMPALPRAVMLIIV
SIFAKKVFLEERKHAETVDLETOFSOSHAKPAQASFTLASVSQLMNDWDDP
VFADKRLGDFALVPADGKYAPAGTAVROLAKRHSIVLENENGVLVLIHLGLTVK
LNGTEFVSVEEGSOVEAGQOILEFMDPAIKQAKLDDTVITVINSSETFANSQMLPI
GHSVALDVEFKEGKN"
1877..4519
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAC07488.1"
/db_xref="GI:9996773"
/translation="MSNKLVEKRRVDQADLAWLDPPEVEYENTIPSHDESFSQOE
LEEGKSLVOSLDQMLIDYAEENGQVNVYAEFDNSNFYSVPNLEFQGGQO
VNVQYPMDSSEETFPPOIRKPNLAIVYVPEDEKEMPEKYSILDKGAATVYML
NGHFVGTGDESFSEPMVTKFLKKEKNRLANVALIKYSSASMLEDDQFWRSGILFRSY
TLOAKPRHLDELKLTASLTDNOKKLEAVANLYLPNASEFLVRSDEGDLVAK
LGPILRSQLEFETLADLPVAMSAKPMUYOYRLIYQASLSEVRSQEGRNELDK
GIMYLNGORIVFKGARNRHEPDSKGRALTEEDMIWDIKTKRSLNINARCSHYNOSL
FYRLCKYGLYVIDEANLESHGTWEKVGHEDEPFENVQDOHMLGASLSVKMMAR
DKNHSALTIWISIGNESYAGTVFQADYVARKADPTROYHGGTGHNRKFDPAATIESR
MYAPAKYIEEYLTNRKPAKPFISYEYAHAMGNSVDDLAAITALEKYPHYOGFTIMNDID
QGLEKDHLLTGDEDFDPTDYFCNGLVPADTESPRLANVALIANKLEVEKQO
LFLKNDULFTNSSSYFLTSLVDGKLTYSRPLTGLEPESGTFALPMPVEVADERG
EYVYRVTAHLKEDLPADDEGTVAAEAEVOKLPDEFKPEGPDVDDYDNLGLGNF
OILFESKVGWPSLKYAGREYLRKLPETFERALITDMDRGAGYDYDIAWMENAKYAR
LKDISCEVKEDSVLVKTAFTLPVALKGDLTVTYTVVDGRKLAIVADPFAEAGLLPA
FGNLALPKELTDVRYGLGPNES"

BASE COUNT 1229 a 1341 c 1323 g 1166 t
ORIGIN

Query Match 43.1%; Score 3497.2; DB 6; Length 5059;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 3518; Conservative 0; Mismatches 13; Indels 3; Gaps 1;
QY 4584 ccgataaaagttagcgacggttgcctgtgtgcccagcagcagttaggtctacgcgc 4643
DB 1479 CCGACAAAAGTTAGGCGACGGCTTGCCTGTGTCCAGCAGCGTAAGCTACGGCGC 1538
QY 4644 cattgcgcgtactgtccgcagcgtgcccagaagccggcaactgcgtcctctgaagaatg 4703
DB 1539 CATTGGCCGTACTGTCCGCCACGCTGGCCAAAGACCCGGCACCTGATCTCCCGGAAATG 1598
QY 4704 aacctggaggtctgtgcttgcattcaaccttggccttgggcagcagttcaaatgaacggagctg 4763
DB 1599 AACATGGGGTCTTGGTCTTTGATTACCTTGGCTGGGACGCTCAATTAAGGGGACTG 1658
QY 4764 gctttgtcaagtatgtltaagaagggcagccaggtagaaagccggcagcagatcctcggat 4823
DB 1659 GCTTTGTCAAGTATGTAAGAAGGCGACCGACGAGTGAAGAGCGCGGCGACGAGATCCTGGAAT 1718

OY 4824 tctggaccgcgcgtacgaagcagccaaagcttgcagcagcagttacgttgaccgtcatca 4883
DB 1719 TCTGGGACCCGGCATCAAGCAGCGCAAGCTGGAGACACGGTAACTGACCGCATCA 1778
OY 4884 acaagcgaacttccaaatagcagatgtccttgcgcgtacgtgcgcagcgtlccaaagcc 4943
DB 1779 ACAGCGAACTTTCGCAATATAGCCAGATGCTCTTCCATGCGCCACAGCTCCCAAGCC 1838
OY 4944 tggatgatgttctcaagttaagaagggaagttaagaataatgaatgaatgaatgaatga 5003
DB 1839 TGGATGATGTATTCAGTTAGAGGGAAGATTAAGAAATAGCAATATAGTAAAG 1898
OY 5004 aaaaaaggttgcacagcagcagccttgccttgcctgcagcagcagcagtttcaagaatca 5063
DB 1899 AAAAAAGGTTGACCGACGACAGCTTGCGCTGGCTGACTGACCGGAACTTTACGAAGTCA 1958
OY 5064 atcaattcccccgcagctcgcagcaatgagtccttccaaagccaggaagaacttgaagag 5123
DB 1959 ATCAATTCCCCCGCACTCCGACATGAGTCTTCCAAAGCCAGCAACTGGAGGAG 2018
OY 5124 gcaagtcgaattagtgtagtccttcgcgcgggagcttgcctgattgactgaactgaag 5183
DB 2019 GCAGTCCAGTTTACTGATGCTCCTGACGGGAGCTGCTGATTGACTGACGTAAGAC 2078
OY 5184 gccagggaccagctcaacttctatgcagaagaacttgcagatagaactttaagtcagta 5243
DB 2079 GCCAGGACACAGTCAACTTCTATGCAGAAAGACTTTGACGATAGCAATTTAAGTCAGTCA 2138
OY 5244 aagtaaccgcgcagccttgcagcagcagccttgcgcagcccccagatgtacagcagc 5303
DB 2139 AAGTACCCGGCACTCGAAGCTGCAAGGCTTTGGCCACCCCAAGTATGCAACGTCAT 2198
OY 5304 atccatggagcgcgcagtagagagatlttccgcgcccacaaatccaaagcaaaatccgcctg 5363
DB 2199 ATCCATGGAGGAGCGAGTGAGGAGATTTCCGCCCAATTCGAAGCAAAATCCGCTCG 2258
OY 5364 cttcttaagcagataacttgaactgtgaatgaagcttcttgggacaaggaagtcaagcttga 5423
DB 2259 CTTCTTAAGTCAAGATACCTTGAACCTGATGAAAGCTTTTGAGACAAAGTCACTTGA 2318
OY 5424 agtttgaagggcgcgcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5483
DB 2319 AGTTTGAAGGGGCGGCAACAGCATCTATGCTGGCTGACAGCGCACCTTGGCGGCTACG 2378
OY 5484 gggagagactccttaccacacagcagcagcagcagcagcagcagcagcagcagcagcagc 5543
DB 2379 GGGAGAGCTCTTACCCCAAGCGAGTTATGTTACCAAGTTCTCTCAAGAAAGAAATA 2438
OY 5544 accgcctgcagctgtccttctcaagaagtattcttcgcgctcctgcgtggaagcagagact 5603
DB 2439 ACCGCTGCGAGTGGCTCTCAAGATATCTTCCGCTCTGCTGAGGACACAGGACT 2498
OY 5604 tctggcagcagctgtcgttcttgcagatcagtgacttccagcagcagcagcagcagcagc 5663
DB 2499 TCTGGCGCATGTCTGTTGTTGATCAATCAGTACCTTCAAGCGCAAGCGCGTGTGACT 2558
OY 5664 tggagagccttaagccttaccacagcagcagcagcagcagcagcagcagcagcagcagcagc 5723
DB 2559 TGGAGAGACTTAACTTACGGCGACAGCTTACCGGATTAACCAAAAGAAAGGCTGGAG 2618
OY 5724 tcgaagccaatatcttccacagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5783
DB 2619 TCGAAGCCAAATATTCCTTACCGCTTCCCAAAAGCCAGCTTAAAGCTGGAAGGCGGGATA 2678
OY 5784 gtaagagtgactgttcttgcgaagaagcctgggcccacatagaagcagcagcagcagcagc 5843
DB 2679 GTGAAGTGACTTGTCTGTAAGAGCTGGGCGCAATCAGAACGACGACGAGTGAATCA 2738
OY 5844 ctctgctgaatttgcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 5903
DB 2739 CTCTGGCTGATTTGGCAGTAGCTGCTGGAGGCGGAAAGGCTTAACCTTACAGGTCC 2798

5904 gccctatattatccagcgagccctcttagaggttagccgaggaagtggttccc 5963
|||||
Db 2799 GCCGTATTATACAGGAGGAGGAGCTCTTAGAGGTTAGCCGGCAGGAATGGGTTTTCC 2858
5964 gcaacttgaactaaagacggygattatgtacctaagcgagcgagatcgtctcaag 6023
|||||
Db 2859 GCAACTTGAACCTAAAGACGGGATTTATGTACTTAAACGGCCAGCGGATCGTCTTCAAG 2918
6024 gggcgaaacggcgagaaatttgacagtaagttgggtcggtcatcacggaagaagatatga 6083
|||||
Db 2919 GGGGCAACCGGACGAAATTTGTACAGTAAAGTTGGCGGGCTATCACAGAAAGAGATTAAGA 2978
6084 tcttgagcatcaagacatgaagcgaaagcaacatcgaatgtctgcgtcgctctacacc 6143
|||||
Db 2979 TCTGGGATATCAAGACCATGAGCGAAGCAATCATGCTGCTCGCTCTCACTAC 3038
6144 cgaacagtcctctcttctacggctctgtgaacaagtcagcgctctacgtcatatgatag 6203
|||||
Db 3039 CGAACCAGTCCCTCTTTACCGGCTGTGTGACAAGTACGGGCTTTACGTCATTTGATGAG 3098
6204 cttaacctggaagcgagccactgggaaaaagtgggggggcgacgaagaatccttagcttca 6263
|||||
Db 3099 CTAACTGGAAGGCGACAGGCACTGGGAAAAAGTGGGGGGCAGCAAGATCTAGCTTCA 3158
6264 atgttcaggcgatgacacagcatgtgtgagcgacgttatccgggtgaagaacatga 6323
|||||
Db 3159 ATGTTCCAGGCGATGACACAGATTTGGCTGGAGGCGACGCTTATCCGGGTGAAGACATGA 3218
6324 tggctcggaagaaacacatgtctcaatcctaactcgtctttaggaatgagctttag 6383
|||||
Db 3219 TGGCTCGGAGCAAGAACCATGCTTCAATCCGATCTGCTGTTTAGCAATAGTCTTACG 3278
6384 ccggactgtcttggccaaatgagctatagctcggaagcgagtcagccggcggttc 6443
|||||
Db 3279 CCGGACTGTCTTTGGCCAAATGAGCTGATTTAGCTCGGAGGCTGATCCGCGGGTTC 3338
6444 agcacta tgaaggggtgacccacaaccggaagtttgaagcgacacacagattgaagcc 6503
|||||
Db 3339 AGCACTGTGAAGGGGTGACCCACAAACGGAAGTTTGAAGGAGCCAGCCAGTTGAAGGCC 3398
6504 ggaatgtatgtctcgagcaaggttaattgaagaatacttgacaaataaacacagccaagcat 6563
|||||
Db 3399 GGATGTGTGCTCGGCGCAAGGTAAATGAAGATCTTGACCAATTAACCAACCAAGCCAT 3458
6564 ttatcctagtgtgaatcagctcagccacatgggcaactcgtctggtgacttggccactaca 6623
|||||
Db 3459 TTATCTAGTTGAATGAGCTACGCGCATGCGCAATCTCGTGGTGACTGGCCGCCCTTACA 3518
6624 cggcccttgaaaaaatacccccaactacacagggcggtcatcctcgtgagacttgaatga 6683
|||||
Db 3519 CGGCGCTGGAATAATACCCCACTACACAGGGCGGCTTCATCTGGGACTGGATTGACCAAG 3578
6684 gactggaaaaagacgggacactgtcttataggggcgagactcagatgacggcgccaacgaat 6743
|||||
Db 3579 GACTGGAAAAAGACGGGACCTGCTTTATGGGGCGACTTGATGACCGGGCAACCGACT 3638
6744 atgaatctcgaggaacggcctggtcttctgtgacggagactgaatcgcgaactgtgcta 6803
|||||
Db 3639 ATGATTAATCTCGGGAAGGCGCTGCTTTGCTGACCGGACTGAATCGCCCAAACTGGCTA 3688
6804 atgtcaagcccttaccgcaacctaagttagaagtaaaagatlyggcagcttctctca 6863
|||||
Db 3699 ATGCAAGGCGCCCTTACGCCAACCTTAAGTTAGAAAGTAAAGATGGGACAGCTTCCCA 3758
6864 aaaaagcaatttattacaaacagctcatctactactctcttgaactagttcttggctg 6923
|||||
Db 3759 AAAACGCAATTTATTTACCAACAGCTCATCTTACTACTCTTGTAGTCTTTTGGTGC 3818
6924 atggaagttgacacacagcgcgcttgacacttggcgtggagccttggcgaaatcgg 6983
|||||
Db 3819 ATGGCAAGTTGACCTACAGAGCGGCGCTCTGACCTTTGGCGCTTGAGACCTGGCAATCCG 3878
6984 ggaacttgccttgcttgccggaagtcgctga tgaaaaaagggaggtcgtctacggg 7043
|||||

3879 GGACTTTGGCCTTGCTTGCGCGGAAGTCCGTGATGAATAAAGAGAGAGTCCGTACCGGG 3938
7044 taacggccacttaaaagaagacttgccttggcgagatgaggttcaactgttgcgtgaag 7103
|||||
Db 3939 TTAGCGCCCACTTAAATAAAGACTTGTGGCGGATGAGAGGCTTCACTGTGGCTGAAG 3998
7104 cagaagaagtgtctcaaaagctgcggnaattagccggaaggggcgcaagattagttg 7163
|||||
Db 3999 CAGAGAAATGAGCTCAAAAGCTGCGGAATTTAACCCGGAAGGGCGCCAGATTTAGTTG 4058
7164 atccgactacaacccaagcttgaagaaagaaataaattccaaattctctcccaagtcga 7223
|||||
Db 4059 ATTCCGACTTACAACTAGGCGCTGAAGAGAAATTAATTCCTTCTTCCAAAGTCA 4118
7224 agggctgacgggttccctcaagtaatgcgggtaaggaaatacttgaacgagctgcggaat 7283
|||||
Db 4119 AGGCTGCGCGGTTTCCCTCAAGTATGCGGATGAGGAATCTTGAAGCGGCTGCGGAAT 4178
7284 ttactcttgccgggcccctgagcaagacggcgagggagctggttagagctatgactgg 7343
|||||
Db 4179 TTACTTGTGGCGGGCCTTGACGGACAAGACCGGGAGCTGTGATCGGCTATGATGTGG 4238
7344 ccggttggaataatgcgcaagtaatgcgcttgaagacatcaagctgcgagtcgaag 7403
|||||
Db 4239 CCGGTTGGAAATGCGCGCAAGTATGCCCGCTTGAAGACATACAGTCGCAAGTCAAG 4298
7404 aagaactcgttcttgcaagaactgtcctttagcgttgcgttgcgtttaaagggtgattaa 7463
|||||
Db 4299 AAGACTCGGTTTGGTCAAGACTGCTTACGTTGCTGTGCGCTTAAAGGGTGAATTTAA 4358
7464 ccgtgacctatgaagtgaatgaacggggcggaagattgtcgttaacagcttcaatccagcg 7523
|||||
Db 4359 CTGTGACCTATGAAGTGCATGAGACGGGCAAGATGTGCTGTAACACTACTTCCAGGCG 4418
7524 cggagaagactgtgtctcttgcacgcttctgtgaacacttgcgcttgcgaagaactga 7583
|||||
Db 4419 CGGAAGAAAGCGGCTCTTTCGACAGCTTGGCTGTAACCTGGCCCTGCAAAAGAACTGA 4478
7584 ccgattaccgttactatagttcttgagactaatgagagcttaccacgagcgttgaagta 7643
|||||
Db 4479 CCGATTACCGGTACTATGTGTTGGACCTAATGAGACTCAACGACCGCTTGGAAAGTGA 4538
7644 attaccgtgagcatcaccaggaagcggtfaaaaaagaactttagccata---tcgtccg 7700
|||||
Db 4539 ATTACCTGGGATCTACAGGAGGACGGTAAAAAAGAACTTTAGCCCATACCTGGCTCGC 4598
7701 aggaacggcgcaaccggaaggttcgctgtgtacacagcttcttgaatgaagaagcggt 7760
|||||
Db 4599 AGGAACCGGCAACCGGAGCAAGGTTCCGTGTGACCACTCTTGTGATGAAGAGGGGCT 4658
7761 tgaatttaccggccaatggggcagaacttgaactgtctgcgttgcataatctgcggcc 7820
|||||
Db 4659 TGGATTTTACCGGCCAATGGGCAAGCTTGAACCTTGTGCTTTCCTCAATTTCTGCGGCC 4718
7821 aaattgaagcagcgagccacacttctgaactgactcaacaattacacttgggttgaagct 7880
|||||
Db 4719 AAATTTGAAGCGGCGACACGCTTTTGAAGTACTAACAATTTACATTTGGGTTAGAGCT 4778
7881 taagcgccagatgggggttcggcggygatagcctcttggggcgagaaggttccaccggaat 7940
|||||
Db 4779 TAAACGCCCANAGTGGGGGTGCGGGGATGACTCTGCGGGGCAAGAGTCCACCCGGAAT 4838
7941 tctgcctgagatgcctcaaaaagcccgcaagcttgccttgggtatgataagcccttctaa 8000
|||||
Db 4839 TCTGCTTGAGTGTCAAAAAGCCCGCAAGCTCCGCTGTGATTTAGAGCCCTTTTACTTAA 4898
8001 aataatgtactaaatgtacttaacaggaatgaatctttagtaaaagcaaaagagtgagga 8060
|||||
Db 4899 AATTAATGCTCAATTTGACTTAAACAGATGAATAATTTTAAAGTAAAGCAAGGAGTGAAG 4958
8061 agatgcaacagatcagaagaatgccaagcgagcggttgcgttaagcgaggt 8114
|||||

|||||
Db 2491 GGAAGACGTCCTTACCCCAAGGAGTTATGTTACCAAGTCCCTCAAGAAAGAAACA 2550
OY 5544 accgccttggaagtggtcctcctaaagtatcttcgcctccttgctggaagccagact 5603
Db 2551 ACCGCTGGCAGGGGCTCTACAAATATTTCTCCCTCTCCCTGGCTGGAGACAGACT 2610
OY 5604 tctggcgaatcttggtttgttcagatcaagtgactcttcagcgaagccgctcgaact 5663
Db 2611 TCGGGGCAATGCTGTGTTGTTTCAGATCAGTGCCTCTTCAGGGCCAAAGCCGCTGCACT 2670
OY 5664 tggaggaaccttaagcttaagccagacttgaccgaactaaccaaaagaaagactggaag 5723
Db 2671 TGGAGAGACCTTAAGCTTACGGCCAGCTTGACGATTAACCAAAAAAGAAAGCTGGAAG 2730
OY 5724 tgaagccaatattgctcctaccgcttgccaaatgcccagcttaagcttggaagtgcggata 5783
Db 2731 TCGAAGCCCAATATTTGCTACCGCTTCCCAATCCAGCTTTAAGCTGGAAGTCGGGATA 2790
OY 5784 gtgaaggtgaacttggttgctgtaaaagctggcccgaatcagaagcgaagcctggaattca 5843
Db 2791 GTGAAGGTGACTTGTTGCTGGAAGACTGGCCCATCAGAACGAGCACTGGAATTCA 2850
OY 5844 ctctggcgaatttgccagtagctgctggaagcgcggaanaagccctaaccttaacaggtcc 5903
Db 2851 CTCTGGCTGATTTGGCCAGAGCTGCTCGAGCCGGAAGAAAGCCTTAACCTTTACAGGTC 2910
OY 5904 gcttgatattataccaggaagcgaagcctcttaaggttagcgcggaagtgtgttcc 5963
Db 2911 GCTGTATTTATTCAGGCGAGGCGCTTTAAGGTTAGCCGGCAGGAAGTGGGTTTC 2970
OY 5964 gcaactctgaactaaagaacggaattatgtacctaaacgcgcagcagatcgtcttcaag 6023
Db 2971 GCACCTTTGAATTAAGAGCGGATTAATGACTTTAACGGCCAGCGGATGCTTTCAAG 3030
OY 6024 gggccaaacggcgcaaatctgaagtaagtgtgtgcgtatcacggaaagagatatga 6083
Db 3031 GGGCCAAACGGCGCAGATTTGACAGTAAGTGGCGGCTATCAAGAAAGAGATGA 3090
OY 6084 tctggagacatcaagacatgaagcgaagcacaatcgtctgctcgtctcactacc 6143
Db 3091 TCTGGATATCAAGACATGAAGCGAAGCAACATCAATGCTGTCCGCTCTCACTACC 3150
OY 6144 cgaaccagtcctccttaccggctcgtgtgacaagtaacggccttaacgtcatgtgaag 6203
Db 3151 CGAACCAAGTCCCTTTACCGGCTGTGTGACAAATGAGCCCTTAAGTCAATGAGAG 3210
OY 6204 ctaaccttggaagcgaacgcgcgccttggaagaaagtggggggcgaagacatcctagctca 6263
Db 3211 CTAACTGGAAAGCCACGCACTGGGAAAAAGTGGGGGGCAGAAAGATCCTAGCTTCA 3270
OY 6264 atgttcgaagcgaatgaacagcattgctggaagccagcttaaccgggtgaaagacatga 6323
Db 3271 ATGTTCCAGGCGATGACCAACATTTGGCTGGAGCCAGCTTATCCCGGCTGAACATGA 3330
OY 6324 tggctcgggaacaaagacatgcttcaatcctaaatctggtctttaggcaatgagcttaag 6383
Db 3331 TGGCTGGGAGCAAGAACATGCTTCAATCCTGATCTGTGCTTTAAGCAATGAGTCTTAG 3390
OY 6384 ccggacactcttctgcccacaaatgagctgattagctccggaagcgtgatccgaaggttc 6443
Db 3391 CCGGCACTGCTCTTGGCCAAATGCTGATTACCTCGGAAGGCTGATCCGACCCGGGTTTC 3450
OY 6444 agcaactgaaggggtgacccacaacccggaagtgtgaagcgcacacagatgtgaagcc 6503
Db 3451 AGCACTATGAAGGGGTGACCCACCAACCGAAGTTTGAACAGCGCACCCAGATTGAAGCC 3510
OY 6504 ggaatgtaagctccggcgaagtaattgaagaaacttgaccacaaataaacagccaaagcat 6563
Db 3511 GGATGATGCTCCGGCCAAAGGTAATTGAAGAAATCTTGACCAATTAACCGCAAGCAT 3570
OY 6564 ttatctcaattgataagctcaagcgaatgggaacccgtcggtgacccgtggccgactaca 6623
|||||

Db 3571 TTAATCTAGTTGAATACGCTACAGCCATGGGCAACTCCGTGCGTGAACCTGGCCCTACA 3630
OY 6624 cggcctcggaaaataatccccactaccagggcggtctcatcttggaattgataccaag 6683
Db 3631 CGGCGCTGGAAAAATATACCCCCACTTACAGGGCGGCTTCACTGGGATGATGACCAAG 3690
OY 6684 gactggaaaagacgggcacgtcttatatgggggcactctcgatagcagcccaacgact 6743
Db 3691 GACTGGAAAAAGACGGGCACTGCTTTATGGGGCGCACTTCGATGACGGGCCAACGCACT 3750
OY 6744 atgaattctgcyggaagcgcttggtcttgctgacccggaactgaatcgccgaactggtca 6803
Db 3751 ATGAATTTCTCGGGAAGCGGCTGTGCTTGTCTGACCGGACTGAATCCGAAACTGGCTA 3810
OY 6804 atgtcaagcccttaacgcacacttaagttaagaagaagatgggcaactcttctca 6863
Db 3811 ATGTCAAGGCGCTTTACGCCACCTTAAGTTAAGTAAGTAAGATGGGCACTCTCTCTCA 3870
OY 6864 aaaaagcaattatattaccacagctcaatccttaactctcttgactagctctttgtcg 6923
Db 3871 AAACGCAATTTATTTATACACAGCTCATCTTAATCTTCTTGACTGACTTTGGTGC 3930
OY 6924 atggcaagttgacctaacagaagcgcgccttgacccttgccctggaagcctgacgaatccg 6983
Db 3931 ATGGCAAGTTGACCTTACAGAGCGGCGCTGTGACCTTTGGCTGTGAGCCTGGGGAATCCG 3990
OY 6984 ggaaccttgccctgaccttgccggaagtgcgtatgaagaagggaagtgcgtacacggg 7043
Db 3991 GGAACCTTTGCGCTTGTGCTTGGCCGGAATCGCTGATGAAGAAAGAGAGTGTCTACGGG 4050
OY 7044 taacggcccaacttaaaagaaagacttgcttgggcggatgagggccttaacgttgctgaag 7103
Db 4051 TAACGGGCCACTTAAAAAGAAAGACTTGCTTGCGGCGATGAGGGCTTCAGTGTGGCTGAAG 4110
OY 7104 caagaagaagtgaactcaaaagctgcgggaatttaagccggaagggcggaagattagttg 7163
Db 4111 CAGAAGAAGTAGTCAAAACCTCCGGAATTTAAGCCGGAAGGGCGCCAGATTTACTTG 4170
OY 7164 atccgaactacaacctagaagcctgaagaagaaataactccaatctctcttccaaagtc 7223
Db 4171 ATTCCGACTCAACACTAGGCTGGAAGGAATTAATCTCAATTTCTTCTCCAAAGTCA 4230
OY 7224 aaggctggccggttccctcaagtaatgcgttagaggaataacttgaaagcggtgcggaaat 7283
Db 4231 AGGCGTGGCGGCTTCCCTCAAGATAGCCGCTGAGGAATCTTGAAGCGGCTCCGGAAT 4290
OY 7284 ttacctcttgggcgccctgacggaacagacccgggaagctggtttaacgtatagcttg 7343
Db 4291 TTACCTTTGCGCGGCGCTGACGAGCAACGACCGGGAGCTGTGTACGGCTATGATCTGG 4350
OY 7344 cccggtgggaataatgcgggaagtatgcccgcctgaaagacatcaagctggaagtcgaag 7403
Db 4351 CCGGCTGGAAAAATGCCGGCAAGTATGCCGCTTGAAGACATCACTGCGAGGTCAAG 4410
OY 7404 aagaactcgttctgtcaagaactgaccttaacgttgccctgctgcgcttaaaagtgatata 7463
Db 4411 AAGACTCCGCTTTGGTCAAGACTGCCCTTACGTTGCTGTGCCCTTAAAGAGTGATTTAA 4470
OY 7464 ccgtgacatgaagtgcagtgagacggggcgaagattgctgtcaaaagctgaactcccaagg 7523
Db 4471 CCGTGAACCTATGAAGTCGATGAGAGCGGGCAAGATTCGTGTAAACGCTGATTCACAGCG 4530
OY 7524 cggaaagaagtgtgctcttgccagcccttgcttggaaccccgcccgccaaagaactga 7583
Db 4531 CGGAAGAAGCGGCTCTTCCAGCCTTTGGCTTGAACCTAAGCCCTGCCCCAAAGAACTGA 4590
OY 7584 ccgaattaccgctactatggtctggaacctaaatgaagcttaaccagaacccgttggaagta 7643
Db 4591 CCGATTACCGCTACTATGCTGTGGACCTTAATGAGAGCTACCAAGACCGCTTGAAGGTA 4650
OY 7644 attacttggtgacataccacgggagcggttaaaagaactttaaccata---tcgtccgc 7700
Db 4651 ATTACTGGGCATCTACCAAGGAGCGGTAAAAAAAGAACTTTACCCATACCTGCTGCCG 4710
|||||

QY	7701	aggaacaggcgcaaccggagcaaggttcgctcgtgtaccactcttgaatgaataaaggcgagct	7760
Db	4711	AGGAACAGGCGCAACCGGAGCAAGTTCCCTGGTACCACCTCTTGTATATAAAGGGGGCT	4770
QY	7761	tggaatttaccgagccaatttggcgcaagccttgaactctgtcgtcttggccatctcgcgcgc	7820
Db	4771	TGGAATTATACGGCCAAATGGGGCAGACTTGAACCTTGTCGTGTGGCANATTCCTGCCGCC	4830
QY	7821	aaatgaagcagcgagccaacgccttltgaactgactacaataacacttgggttagagcct	7880
Db	4831	AAATGGAAGCACCGGACACACGCTTTTGAACGTACATAAATTAACCTTGGGTTAGAGCT	4890
QY	7881	taagcgccacagatggggggtcggcgcgagatgcctcttggggcgagaaggtccaccggagat	7940
Db	4891	TAAAGCGCCACAGATGGGGGTCGGCGCGGAGATGACTCCTGGGGCGAGAGGTCACCGGAAAT	4950
QY	7941	tctgcctgtagtgcataaaagccgcgcagcttcgctcgtgtatcaagcccttactaa	8000
Db	4951	TCTGCTCGATGCTCTAAAAAGCCCCCGACGCTCGGCTGTGATTCAGCCCTTTACTAA	5010
QY	8001	aataa	8005
Db	5011	AATAA	5015
RESULT	4		
LOCUS	EVE132038		
DEFINITION	Expression vector pCDP1.	3840 bp	MRNA
ACCESSION	EVE132038		
VERSION	AJ132038.1	GI:5640088	
KEYWORDS	AMP gene; beta lactamase; ColE1 origin of replication; multiple cloning site; Sp6 promoter; SV40 origin of replication; T7 promoter; xanthine-guanine phosphoribosyl transferase; Xanthine-guanine phosphoribosyl transferase gene.		
SOURCE	Expression vector pCDP1.		
ORGANISM	Expression vector pCDP1.		
REFERENCE	1 (bases 1 to 3840)		
AUTHORS	Zeng,B.J.		
TITLE	Mammalian Expression Vector for with fuse Xanthine-guanine phosphoribosyl transferase Tag		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3840)		
AUTHORS	Zeng,B.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-FEB-1999) Zeng B.J., Gene Engineering Center, Institute of Microbiology, Zhongguancun, Beijing, Beijing 100080, CHINA		
FEATURES			
source	Location/Qualifiers		
	1..3840		
	/organism="Expression vector pCDP1"		
	/db_xref="taxon:90749"		
	209..863		
	/note="CMV"		
promoter	864..882		
	/note="T7"		
misc-feature	882..984		
	/note="Multiple cloning site; HindIII, BamHI, BclXI, EcoRI, NotI, XhoI"		
CDS	929..1387		
	/codon_start=1		
	/product="Xanthine-guanine phosphoribosyl transferase"		
	/protein_id="CAB51567.1"		
	/db_xref="GI:5640089"		
	/translation="MSERYIYTWDMLOIHAKLASRLMSEQWKGIIAIVSRGLVPGA		
	LIARELGIHVADVICISSYHDNORELKVLRAGGDEGPIVIDLVDTGTAVAIRE		
	MYPAHVITIPAKPAGRPVLDVVDYIPDPTWIRQPDMDGVFVPPISGR"		
	1649..1863		
	/note="BGH"		
	2450..2775		
	/note="Sp6"		
polyA_site			
promoter			

rep_origin	2644..:2729
gene	/note="SV40"
gene	complement (2844..:3704)
CDS	/gene="amp"
	complement (2844..:3704)
	/gene="amp"
	/codon_start=1
	/product="beta-lactamase"
	/protein_id="CAB51568.1"
	/db_xref="GI:5640090"
rep_origin	3632..>3840
	/note="COLE1"
BASE COUNT	938 a 960 c 964 g 978 t
ORIGIN	
Query Match	16.5%; Score 1336.4; DB 12; Length 3840;
Best Local Similarity	75.4%; Pred. No. 1.9e-306;
Matches 1991: Conservative	0; Mismatches 16; Indels 635; Gaps
1	gatgacggcgccaatatagcgtgtgacatgtatattgacacgtttaatagtaatca 60
135	GATGACGGCGCCAGATATACCGCTTGACATGATTATGACTAGTATTAATGATATCA 194
61	attacgggtcatatagttcctaagcccatatagtgaagttccggttataactaactagtgta 120
195	ATTACGGGTGCTATTAGTTCATAGCCCATATATGAGATGTTCCGCTTACATTAATTAACGGTA 254
121	aatggccgcgctgtgcgtgacgcccacaagacccccgcccatgtgacgctcaataatagcgtatc 180
255	AATGGCCGCGCTGTGCTGACCGCCCAACGACCCCGCCCATTTGACGTAAATTAAGAGGTAT 314
181	gttcccatagtaacgccaatagggaacttccatgtgacgtcaatagggtgacatatttaacgg 240
315	GTTCCCATAGTAACGCCCATAGGGACATTTCCATTGACGTCATAGGGGGACATTTTAACGG 374
241	taaatgtcccaattgtggacagtaacaaagtgtacatatgccaagttaagcccccatatgac 300
375	TAAACTGCCCATTTGGCGAGTACATCAAGTGTATATATGCCAAGTACGCCCTATTATGAC 434
301	gtcaatgacggtlaaatggtccgcgtgtgcataatgcccagtaacatgaccttaaggacctt 360
435	GTCATATGACGGTAAATATGGCCCGCTGCAATATATCCAGTACATGACCTTATGGACTTT 494
361	cctaattggcaggtacatcattacgtatttgtaacgtcttatccaaagtgtacgtggcttttg 420
495	CTTACTTTGGCGATACATTTAGCTTTATGCTATTCCTCTTTACCATGATGATGCCGTTTGG 554
421	cagtaacatcaaatggtgcgtgtgaatacgggttttgacacacggggagatttccaagttctcaccc 480
555	CAGTACATCAATATGGCGCGGTATACGCTTTAGCTACAGGGGATTTTCCAAGTCTCCACCCC 614
481	attgacgtcgaatggtgagttgttttggtcaccaaaaatacaagggaacttccaanaatgtcgt 540
615	ATTGACGTCAATGGAGATTTTGTTTGGACACCAAAATATACGGGACCTTCCAAATATGTCGT 674
541	aacaactccggcccaattgaacgaatggcggtatggcggttgaacggttgaaggtgactatata 600
675	AACAACCTCCGCCCATTTAGCCCAATATGGCGCGGTATGGCGGTATACGCGTGGAGGCTATATA 734
601	agcagaagctctctgtgctaactagaaaccaactgtctactggtcttaacgaataatbaacg 660
735	AGCAGAAGCTCTCTGCTTAACATAGAAACCCACTGCTTAAGTGCTTAACGAAATTAATACG 794
661	actaactataaggagaccacaagcttgttaccgagctgtgataccaactgtaacggccggca 720
795	ACTCACTATTAAGGAGAACCAAGCTTGTGATACGAGCTGTGATCCATATGATTAACGGCCGCA 854

Oy	721	gtgtgctcggaattcttgaagatatacaatacaacacgaggcgccgtctgagaatgcataga-	773
Db	855	gttgtctggaatttctgcagatattccatcacactggcggccgtctgcagcatgcatctaaag	914
Oy	780	-----	779
Db	915	ctgggacactttcacatgatgagcaaaaattacatgtctacgtgggacatgttgagatccatg	974
Oy	780	-----	779
Db	975	cacgtaaacctcgcaagccgacgtatgcttgcgaacaaatgaaaggacattattggcgtaa	1034
Oy	780	-----	779
Db	1035	gccctggcggctgtgctacccgggtgcgttactggcgcgtcactgggtattcgtcatgtgc	1094
Oy	780	-----	779
Db	1095	ataccgtttgtattttccagctacgacatcagcaaacacagcgagcttaagtgctgaac	1154
Oy	780	-----	779
Db	1155	gcgcagaagcgatggcgaaagccttcattcgtttattgatcctggtgcgtaccgggtgta	1214
Oy	780	-----	779
Db	1215	ctggcgttgcgattcgtgaatgtatccaaaagcgacattgttcacatcttcgcamaac	1274
Oy	780	-----	779
Db	1275	cgcgtgctcgtccgctgctggtgatgactangttgttgatgtcccgcaagatfaccgattg	1334
Oy	780	-----	779
Db	1335	aacagccgtgggaatattggcgctgctatgtctcccgcaattccgctgcctaatcttttc	1394
Oy	780	-----	779
Db	1395	aacgcctggacactgcgcggcggtgttcttttaacttcaggcggtttacaaatagtcca	1454
Oy	780	-----	779
Db	1455	gtaaagtattctggaagcctgcattcatgacacagggcaaaccttagcgaaacctgttc	1514
Oy	780	-----	805
Db	1515	ccccgttttaaacattcctgaaacctcgacgcctgagggcgccctattctatattgtccat	1574
Oy	806	atgttagagcttgctgatacgaactgacgtgctctctagtctgcaagcactgtgtgtt	865
Db	1575	atgtctaaagctcgtgtgatgacctgcacgtgtgcctttctgtgttcacagccattcgtgtt	1634
Oy	866	gccccctcccgctgacctctccttgaaacctggaagtgccacatcccaactgctcttcctat	925
Db	1635	gccctctcccccgtgccttcttacccttgaaagctccacactgccactgctcttctctaat	1694
Oy	926	aaaatgagaatgcatcgacatctgcctgagtagtgatcatctcatctctgagggtggg	985
Db	1695	aaaattgagaaattgcatctccattgtctgtagtgatgtcatattcttctggggcggtggg	1754
Oy	986	tggagcagacagaacaaggggaagattgggaagacaatagcagcatctgggaatggg	1045
Db	1755	tggggcagcagcagcaaggggaagattgggaacacatgcaagcaatgctggggatggcg	1814
Oy	1046	tgggctctatgctctctgaagcggaaagaacacagctgcatctaaatgaatcgagcaacg	1105
Db	1815	tgggctctatgctctctgaagcggaaagaacacagctgcatctaaatgaatcgagcaacg	1874
Oy	1106	ggggagaagcggtttctgcatcttgaggcctctccgcttctctgctcactgtaactgctg	1155
Db	1875	ggggagaagcggtttctgcatcttgaggcctctccgcttctctgctcactgtaactgctg	1934

OY	1166	ctggctgcttgcgcgcgcgcgcgcgttatcagctcaatccaaaggcggttaac	1225
Db	1935	cttcggttcgttcgcttgccgcgcgcgtatcacctactctaagaagggtatracggtttac	1994
OY	1226	cacgaatcacgggataaacgcaggaagaacaatgtgagccaagaagccagcaaaagccag	1285
Db	1995	cacgaatatcaggcgattAACGCAGGAAGAACAATGTGAGCAAAAGGCCAGCAAAGGCCAG	2054
OY	1286	gaaccgtaaaaaagccgcgcgttgtcgtgtttttccataagctccgcccccctgacagaca	1345
Db	2055	GAAACCGTAAAAAAGCCGGCTGCTGCTGGCTTTTTCATGAGGCTCCCCCCCCTGACGAGACA	2114
OY	1346	tccaaaatatcgaacctaaafcaagaagtgtggaaaacccgaagagcataaagaatacca	1405
Db	2115	TCCAAAATAATGACCCTCAAGTCAAGAGTGCGGAACCCGACAGACATTAAAGATPACA	2174
OY	1406	gacgtttcccccttgtaaagctccctcgtlbgcctctcctgtctccagaccctgacgttacgg	1465
Db	2175	GAGCGTTTCCCCCTGGAAGCTCCTCTGTGCGCTCTCTGTTCGGACCCCTGCGCTTACGGG	2234
OY	1466	ataacctgcgcgcctttccctccttcgggaagcgttgacgtttctcaatgctcagctgttag	1525
Db	2235	ATACCTGTCCGCGCTTTCCTCCCTTCGGGAAGCCTGGCGCTTCTCATATGCTCACGCTGTAG	2294
OY	1526	gtatctaatgttcggtgataggtgtcgttcgcgtccaaagcttgagcgtgtgtgtgcaacgaaccccccg	1585
Db	2295	GTATCTCAAGTTCGGGTGATAGGTGTCGTTCGCTCCAAGCTGGGGCTGTGTGCAGAACCCCCCGT	2354
OY	1586	tcaagccgaacgcgtlbgcgccttalcgcggtfaactatcgctcttgatgcaccaaccggtlaagaca	1645
Db	2355	TCAGCCCGACGCGTCCGGCTTATCCGGTAACATATGCTGTGTAGTCCAACCCGCTTAAGACA	2414
OY	1646	cgaactatcgcacatgtgcagcaagcacatgytaacaggaatlagcagacgaggtatgtag	1705
Db	2415	CGACTTATCGCCACTGCGCAGCAGCCACTGTGTMAACAGATTATACAGAGCGAGGTATGTAGG	2474
OY	1706	cgggtcctacagaagttcttgaaatgtgtgacctaacagagctacacatagaaagagattt	1765
Db	2475	CGGTCTACAGAGTCTTGTGAAGTGGTGGCTTACCTACGCGTACACTAGAGGACAGTATT	2534
OY	1766	tgtatctgcgcgtctgcgtgaagccagttaccttcggaaaaagatgttgtagctcttgatc	1825
Db	2535	TGGTATCTCGCGCTCTGCTGGAACCCAGTTACCTTCGGAAAAAGATTGGTAGCTCTTGATTC	2594
OY	1826	cggcaacaacaaccacgcgctgtgtagcgtgtgttttttcttgcttgcaagcagcagatlacg	1885
Db	2595	CGCGAAACAAACACACCGCTGTGAGCGGTGGTGTTTTGTGTTTTCAGCAGACAGATTAACGGG	2654
OY	1886	cagaaaaaaaagatctcaagaagaatcccttgatccttctacagyggtctgaagctcagtg	1945
Db	2655	CAGAAAAAAAGATCTCAAGAAGATCCCTTGTCTTTTCTACGGGGCTGTGACGCTCAAGTG	2714
OY	1946	gaacgaanaacacagcttaagagatttgctacatgagcgsatcatatatgaatgfatatta	2005
Db	2715	GAAACCAAACTCACGTTTAAGGATTTTGTGCTCATGAGATTATCAAAAAGATCTTACCTTA	2774
OY	2006	ga 2007	
Db	2775	GA 2776	
RESULT	5		
AXI07924			
LOCUS	AXI07924	11592 bp	DNA linear PAT 30-APR-2001
DEFINITION	Sequence 52 from Patent W00125466.		
ACCESSION	AXI07924		
VERSION	AXI07924.1	GI:13923307	
KEYWORDS			
SOURCE	synthetic construct.		
ORGANISM	synthetic construct		
REFERENCE	1 (bases 1 to 11592)		
AUTHORS	Slingsby,J., Kingsman,S.M., Rohll,J.O. and Slade,A.O.		

TITLE Producer cell for the production of retroviral vectors
JOURNAL Patent: WO 0125466-A 52 12-APR-2001;
Oxford Biomedica (UK) Limited (GB)

FEATURES
Source 1. 11592
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Synthetic sequence"

BASE COUNT 3058 a 2691 c 2919 g 2924 t
ORIGIN

Query Match 16.1%; Score 1307.4; DB 6; Length 11592;
Best Local Similarity 81.5%; Pred. No. 1.8e-299;
Matches 1618; Conservative 0; Mismatches 216; Indels 151; Gaps 3;

Qy 23 gttaacattgatttgactagtatttaataagtaatacaatcaggagcattatcata 82
Db 8336 GTTACATTGATTAATGACGTATTAATTAATCAATTAACGGGGCATTAAGTCTCA 8395
Qy 83 gccacataatgagctccgcttaacatacgtataatgagccgcctgctgacgc 142
Db 8396 GCCCATATATGAGATTCGGCGTTACATTAAGTAAATGGCCCGCTGGTACCGC 8455
Qy 143 ccaacgacccccccacattgacgtcaataatgacgtatgctcccaatgaaagccaaatag 202
Db 8456 CCAACGACCCCCCGCATTTGACGTCAATTAATGAGTATGCCATTAATACGCCCAATAG 8515
Qy 203 ggaattccattgacgtcaataatgaggttgacatttaacgttaacatgcccacttgacagac 262
Db 8516 GGACTTTTCATTGACGTCAATGAGGTGAGTATTTAAGGTAAACGCCCACTTGGCAGTAC 8575
Qy 263 atcaagtgatcatatgacgaagtaacgccccattgacgtcaataatgacgtataatgagccg 322
Db 8576 ATCAAGTATCATATGACCAAGTACGAGCCCTATGACGTCAATGACGTAAATGGCCCG 8635
Qy 323 cctggcaataagccaggaatgacattatgaggaattccacttctgacgtacgtacatcacg 382
Db 8636 CCTGGCATTAAGCCAGTACATGACCTTATGAGACTTCTACTTGGCAGTACATCTACG 8695
Qy 383 tatagtcacgtatatacctatgtagtgcgttttgacgtatatacgaatgagcgtagat 442
Db 8696 TATTAGTCATGCTATTAATCAATGATGATGCGTTTGGCAGTACATCAATGGCGCTGAT 8755
Qy 443 agcggtttgacacagggatltccaaagtcaccacacgtgaagtcgaatggagtttgt 502
Db 8756 AGCGGTTGACCTACGGGGATTTCCAAAGTCCACCCCATTTGACATGAGGAGTTTGT 8815
Qy 503 tttagacacaaatcaacggaggaatttccaaatgtcgttaaacacacgccccattgagcg 562
Db 8816 TTGGACACCAAAATCAACGGGACTTTCCAAATGTCGTAAACAACGCCGCCATTGACGC 8875
Qy 563 aaatgggagtagcggtgtagcggtgtaggtctatataagcagagactctctggtcaacta 622
Db 8876 AAATGGGGGTAGGAGCATGTACGGTGGAGGTCTATATAAGCAGACCTGTTTATGGAAC 8935
Qy 623 gagaacccactgctactgcttatcgaaataataacgactcaactataggagaccgaag 682
Db 8936 GGGGACCTCAGATTCTGCGGTCTGAGTCCCTCTCTGCG-----TGGGCTGAAAGAGC 8986
Qy 683 ctgttgacagagctcgtagtcaactagtaacggccgcaggtgtgctgagatcttgagata 742
Db 8987 CTTTGATTAATAATTAATTTCTACTACGTCCCTGCTCTCTAGTTGTCTGTTGAGATCC 9046
Qy 743 tccatcaactgagcgccgctcgagacgtacgtactagagggccctaattcatagtgtcaac 802
Db 9047 TACAGAGCTCATGCTTGGCGTAATCATGTCATAG----- 9082
Qy 803 taatgtcagagctcgctgtagcagcctcagactgtgctctcctagttgcccagcactgtgtg 862
Db 9083 -----CTGTTTCTGTTGCTGAATTTG 9102
Qy 863 tttagccctcccccgtgctctccttgacacctggaaggtgccactcccaactgtctcttcc 922

Db 9103 TTATCCGCTCACAATTCACACAAATACAGACCGGAGCATTAAGTGAAGCTCGGG 9162
Qy 923 aataaataaggaatgtgacatgcattgtctgagtagtgctcatctctatcttgagggtg 982
Db 9163 TGCCTAATGATGAGCTAACTACATTAATTAATGTCGTTGGCTACCTGCCCGCTTCCACTC 9222
Qy 983 ggtgtgggacagacagcaaggaggagatttggaagaacatagacgactgtggagatg 1042
Db 9223 GG----- 9224
Qy 1043 cgttggcctctatgctctcgtgagcgggaagaacacagctcattatgaatcggccaagc 1102
Db 9225 -----GAAACCTGTCCTGCGACGCTGATTAATTAATCGGCAACG 9264
Qy 1103 cgcgggagagagcggtttgctgattggcgctcttcgcctccctccgctcactgactgct 1162
Db 9265 CGCGGGAGAGCGCGTTGGCTATTTGGCGCTCTCCCTTCCTCCCTGCTACGTACTGCTCCT 9324
Qy 1163 ggcgtcgtctcgtcgtcgtcgagcgtatcaactcaactcaaaagcgtataatgct 1222
Db 9325 GCGCTCGGTGCTGCGCTGCGCGAGCGGTATACGCTCACTCAAAAGCGGTAAATACGCTT 9384
Qy 1223 atccacgaatcaggggtatcaacgcaggaagaacatgtgagcgaagaagccagcaaaagc 1282
Db 9385 ATCCACAGAAATCAGGGATACGACGAGAAAGCAATGTGACCAAAAGGCCAGCAAAAGGC 9444
Qy 1283 caggacccgttaaaaagcgcgctgtcgtgcttttccatagagctcgcgccctcgaaga 1342
Db 9445 CAGGAACCGTAAAAAGCGCGCTTCTGCTGCTTTTCCATAGGCTCGCCCGCTCAGAGA 9504
Qy 1343 gcatcacaanaatcgacgcctcaagtcagaggttgcggaaccgcagacactaataagata 1402
Db 9505 GCATCACAANAATCAGCCCTCAAGTCAAGGTGCGGAACCCGAGACATATAAAGATA 9564
Qy 1403 ccaagcgcttcccccctggaagctccctcgtcgtcctcctctcttcgcagccctgcgctac 1462
Db 9565 CCAAGCGTTCCCTCGAAGCTCCTCTGCGCTCTCTGTTCCGACCTGCGCTTAC 9624
Qy 1463 cggatacctgtccgccttctccctccttcggaagcggtgagccttccaaatgctcaagctg 1522
Db 9625 CGGATACCTGTCGCCCTTCTCCCTCGGAAGCCTGCGCGCTTCTCATGCTCAGCTG 9684
Qy 1523 taggtatctcaggttcgtggtgtagtgcgtcccaagctggagctgtgagcagaaacccc 1582
Db 9685 TAGGTATCTCAGTTGGTGTAGGTGCTGCTCCAAAGTGGGCTGTGTGCAAGAACCCCTC 9744
Qy 1583 cgttcagccgacgcctgcgccttataccgtaactacgtcttgatgcacacccgataag 1642
Db 9745 CGTTAGCGCCGACCGCTGCGCTTATCCGTTAATGCTGTGAGTCCAAACCCGCTAAG 9804
Qy 1643 acaagactatcgcacactgacagcagcactgtgtaacagagattagcagagcaggtatgt 1702
Db 9805 ACACGACTTAATCGCACACTGAGCGACGACCTGTAAAGAGATTAACACAGCAGAGTATGT 9864
Qy 1703 agcggtgtcacaaggttcttgaagtggtggtcctacacaggtcactacatagaagaagct 1762
Db 9865 AGCGGCTCTACAGAGTTCTTGAAGTGTGGCTAATCAACGAGCTACATAGAGAGACGT 9924
Qy 1763 atttgatctcgtcgtctcgtgtaagccagttacctcgtgaaanaagttgtagtctctg 1822
Db 9925 ATTGTGATCTGCGCTCTGCTTAACCAAGTACCTTCGAAANAAGTTGTTAGCTCTTG 9984
Qy 1823 atcggcaaaaacacacacgcgtgtgtagcggttttttttlttgcaagcagcagatlaac 1882
Db 9985 ATCCGCAAAACAAACACACGCGGTGAGCGGTGTTTGTGTTTCAAGCAGACATTAAC 10044
Qy 1883 gcgcagaaaaaaggatcacaagaagatcccttgatcttctacaggggtctgacgtca 1942
Db 10045 GCGCAGAAAAAAGGATCTCAAGAAAGATCTTGTGATCTTTCTACGGGGGTGAGAGCTCA 10104
Qy 1943 gtggaacaaaactacgtttaaggagatttgtagtagggagatacatattgaatgtcat 2002


```

Db 1710 CTGTGTGAATTTGTTATCCGCTACAAATTCACACACATACGACCCGGAACATTAAGT 1769
Oy 1041 -----tgcggtggtctatag 1057
Db 1770 GTAAAGCTGGGGGTGCTTATGATGAGCTAACTCACTAATTAATTTGCTTGGCTCACTAC 1829
Oy 1058 ctctc-----gaagcgaaagaaacagctgcatatgaatcgagcaacgacgacg 1107
Db 1830 CCGCTTTCACAGTCCGGGAAACCTGTGCTGCACACTCATTAATGAATGCGCAACCGCGCG 1889
Oy 1108 ggaagggcggttggttggttggttggttggttggttggttggttggttggttggttggt 1167
Db 1890 GGAGAGGGGTTTGGCTATGGGCTCTTCCGCTTCCGCTCACTGACTGCTGCTGCT 1949
Oy 1168 cgttcgttcgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1227
Db 1950 CGGTGCTTGGCTGGGGGAGGGGTATCAGCTCACTCAAAAGCGGTAAATACGTTATCA 2009
Oy 1228 cagaatcaggggataacgagaaagaaacatgtgagcaaaagggcagcaaaagggcagga 1287
Db 2010 CAGAAATCAGGGGATTAACGAGCAAGAACATGTGAGCAAAAGCCAGCAAGGCCAGGA 2069
Oy 1288 accgttaaaaggcgcggttggttggttggttggttggttggttggttggttggttggt 1347
Db 2070 ACCGTAAAGGCGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2129
Oy 1348 acaaaatcagcgtcagatcagagatgtgagcaaacgcgcgcgcgcgcgcgcgcgcgcgcgc 1407
Db 2130 ACAAATATGACGCTCAAGTCAAGAGTGGCGAAACCCGACGAGCTATTAAGATACCAAG 2189
Oy 1408 cgttcgcccttggaagctcccgctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1467
Db 2190 CGTTTCCCTTGAAGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2249
Oy 1468 accgttcgcgccttcctccctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1527
Db 2250 ACCGTGCGCCCTTCTCTCTTGGGAGAGGTGGCGCTTCTCAATGCTACCTGTAGT 2309
Oy 1528 atctcagttcgttgtagtgcgtcccaagctggtcgtgtgtgcaagaaaccccgcttc 1587
Db 2310 ATCTCAGTTCCGTTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2369
Oy 1588 agcccgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1647
Db 2370 ACCCGACCGCTGCGCTTATCCGTTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2429
Oy 1648 actatgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1707
Db 2430 ACTTATCCGCACTGGCAGCAGCTGTTAAGAGATTAGCAGACGAGGTATGTAGCG 2489
Oy 1708 gttctacagagttcttgaagtggtggtcctaactagcgcgtacactagaagagacagat 1767
Db 2490 GTGTACAGAGTTTGAAGTGTGAGCTTAACCTACGCGTACACTGAAGAGACAGTATTG 2549
Oy 1768 gttctgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1827
Db 2550 GTATGTGCGCTCTGCTGAAGCAGTACCTTGGAAGAAAGAGTTGTTAGCTTATATCG 2609
Oy 1828 gaaacaacacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1887
Db 2610 GCAAAACAAACACCGCTGTAGCGGTGTTTGTGTTGCAAGACGCGAGATTACGCGGA 2669
Oy 1888 gaaaaaagagatcacaagaagatccttcttcttcttcttcttcttcttcttcttcttctt 1947
Db 2670 GAAAAAAGGATCTCAAGAGATCTTTGATCTTTTCTACGCGGTGCTGACGCTGATGGA 2729
Oy 1948 acgaaactcacgttaagagatttggtcaagcagagataacttgaatgatattaga 2007
Db 2730 ACGAAACTCACGTTAAGGATTTGTGTCATGATTAATCAAAAGAGATCTTCACTATGA 2789

```

```

AR098190
LOCUS AR098190 3853 bp DNA Linear PAT 14-FEB-2001
DEFINITION Sequence 5 from patent US 6074850.
ACCESSION AR098190
VERSION AR098190.1 GI:12807447
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 3853)
Antelman,D., Gregory,R.J. and Willis,K.N.
TITLE Retinoblastoma fusion polypeptides
JOURNAL Patent: US 6074850-A 5 13 JUN 2000;
FEATURES
source
BASE COUNT 936 a 987 c 941 g 989 t
ORIGIN
Query Match 15.3%; Score 1240.4; DB 6; Length 3853;
Best Local Similarity 74.1%; Pred. No. 1.3e-283;
Matches 1911; Conservative 0; Mismatches 96; Indels 573; Gaps 4;
Oy 1 gatgtacggcgacagataacgacgttgatgacattgattgactagtattatagatca 60
Db 210 GATGTACGGCGCCAGATATACGCGTTGACATGATTAATTAATTAATTAATCA 269
Oy 61 attacgggtcattatgattacatagcccatatagagctcgcgttacaacttaacgta 120
Db 270 ATTACGGGTCATTATGTTCAATAGCCCATATATAGAGTTCCGGGTACATTAATTA 329
Oy 121 aatggccgcgtgctgacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Db 330 AATGGCCGCGCTGCTGACCGCCCAACGACCCGCCCATTTGACTCAATATGACTAT 389
Oy 181 gttccatagtaacgcgaatagagacttcattagcgttaaatggttgactattagc 240
Db 390 GTTCCATAGTAACCCCAATAGGACTTCCATTAGCGCAATGAGGTGACTATTAGCG 449
Oy 241 taactgcgcacttgacgatacatcaatgtaataatgtaacgagacgccccattagc 300
Db 450 TTAATGCCCACTTGGCAGTACATCAATGATATGATGCAAGTACGCCCTTATGAC 509
Oy 301 gtcaatgacgtaaatgacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
Db 510 GTCAATGACGCTAATGAGCCGCCCTGCAATTAAGCCAGTACATGACCTTATGGACTTT 569
Oy 361 cctactggcagttacattagattagtaactgctatgacatgagtgatgcggttgg 420
Db 570 CCTACTTGGCATATCTACGTATTAGTCATCGCTATTACATGTAATGCGGTTTGG 629
Oy 421 cagtacatcaatgggcgtgtagcggttgactcaacgaggaattccaaagttccacc 480
Db 630 CAGTACATCAATGGGCGTGGATAGCGGTTTGACTACGCGGATTTCCAAGTCTCACCC 689
Oy 481 attgacgtcaatggagattgttggcaacaaatcaacgagacattccaaatgctgc 540
Db 690 ATTGACGTCAATGGAGTTGTTTGGCACCAAAATCAACGGGACTTCCAAAATGTCGT 749
Oy 541 aacaactcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 600
Db 750 AACCACTCCGCCCAATTACGCAAAATGGCGGTACCGGTGAGGTCTATATTA 809
Oy 601 agcagagctcttgctactactagaaacacactgcttactgcttactgaaataatagc 660
Db 810 ACGACAGCTCTTGGCTATACATAGAACACCACTGCTTACTGCTTATGAAATTAATACG 869
Oy 661 actcaactatggagacccaagctt----- 685
Db 870 ACTCACTATAGGAGACCAACGAGCTTCCGCGGGGTACACACTGCTTCCGATCGCTGCTG 929
Oy 686 ----- 685

```

Db 930 CGAGGCGCAGCTTTGGGCTCGCGTTGAGGACAAACTCTTCGGGCTTTCCAGTACTC 989
 QY 686 -----gltacagagctcgagtcacatagtaagcgccgcaagtgtgc 726
 Db 990 TTGGATCGGAACCCGTCGGCCTCCGAACGGTACTCCGCCACCGAGGAGACTGAGCAGT 1049
 QY 727 tggatctcgagatatacaatacactgagcgccg-----tgag 767
 Db 1050 CCGGATGACCGGATCGGAAAACCTCTCGAGGCGCGCGTGCAGTCTAGACGAATTCGCG 1109
 QY 768 catgcatctagaaggccctatctctatagtlcacaataatgtagagctcgctgacag 827
 Db 1110 TAGCATTCATGATGGGCCCTATTCTATAGTGCACCTAAATGCTATGAGACTGGCTATAGC 1169
 QY 828 ctgactgtgctctctagatgtgcaagcactgltgtgttgcctcccgctgctcct 887
 Db 1170 CTCGACTGTGCTTCTAGTTGCCAGCCATCTGTTGGCCCTCCCGCTTCTTCTT 1229
 QY 888 gaccttgaaagtgccactccactgctcttccctaaatgaggaattgcatgca 947
 Db 1230 GACCCCTGGAAGTGCCTCCACTGCTCTTCTTAATAATGAGAAATTCGATGCA 1289
 QY 948 tctctgaagtagtgcattctatctctgagggtgagggtgaggcaagcaagggga 1007
 Db 1290 TTGCTGAGTAGGTGTCATTTCTATTTGGGGGCTGGGGGAGGACAGCAAGGGGA 1349
 QY 1008 ggaattggaagaacaatgcaagcactgctgagg----- 1040
 Db 1350 GGATTGGGAAGACAAATAGCCGAATGACCAAGCAGCCCAACCTGCATCAGAG 1409
 QY 1041 ----- 1040
 Db 1410 ATTTCATTCACCCGCCGCTTCTATGAAAGTTGGGCTTCGCAATGTTTCGGGACG 1469
 QY 1041 ----- 1040
 Db 1470 CCGGCTGATGATCTCCACGCGGGGATCTCATGCTGAGATTCTTCCGCCCACT 1529
 QY 1041 ----- 1040
 Db 1530 TGTATTGACAGCTTATATGTTTACAAATTAAGCAATAGCATCACAATTTGACAAATA 1589
 QY 1041 ----- 1040
 Db 1590 AAGATTTTTTCACTGATTTAGTGTGGTTTGCCAAACTCATCATGTAATCTTATC 1649
 QY 1041 ----- 1040
 Db 1650 ATGTCTGTATACCGTCGACCTCTAGCTAGAGCTTGGCGTATCATGTGCTAGCTGTTTC 1709
 QY 1041 ----- 1040
 Db 1710 CTGTGTAATTTGTTATCCGCTCACAATTCACACAAATACGACGCCGGAACATTAAGT 1769
 QY 1041 -----tgcggtgagctctatgg 1057
 Db 1770 GTAAGCGCTGGGTCCTATATGATGAGTACTACATTAATTTGCGTTCGCTCATGTC 1829
 QY 1058 ctctc-----gagcggaagaagaacaagctgcaataaataatgcgcaagcgcg 1107
 Db 1830 CCGCTTTCACGTGCGGGAACCTGTCGCGCAGCTGATTAATGAAATGCGCAACGCCGCG 1889
 QY 1108 ggaagagcggttgcgfatctggcgctctccgcttccctcccaactgactgactgtgc 1167
 Db 1890 GGAAGAGCGGTTTGGCATTTGGGGGCTCTTCCGCTTCTGCTACGACTCCGTGGCT 1949
 QY 1168 cggctcgtcgctgagcgagcgatcaactcaactcaaaagcggtlaatacgtttatca 1227
 Db 1950 CGGTCGTTGCGGTGCGGCGAGCGGTATACACTCACTCAAAAGGGGTATATCGGTTATCCA 2009
 QY 1228 caaatacagggataacgcaagaagaacatgtgagcaaaagccagcaaaagccagga 1287

Db 2010 CAGAAATCAGGGGATTAACGACGAAAGAACATGTAGCAAAAAGGCCAGCAAAAGGCCAGGA 2069
 QY 1288 accgttaaaagcgcggttctggtgttlttccataggtccgcgcctcccgagagcatc 1347
 Db 2070 ACCGTAAAAAGCGCGCTTCTGCGCTTTTTCATAGCGTCCGCCCTCCGAGAGATC 2129
 QY 1348 acaaaaatcgactcaagtcagaggttgggaaaccccgagagactataagaatacagg 1407
 Db 2130 ACAAAATTCAGCTCAATGACAGGTGCGAAACCCGACAGGACTTAAATATCCAGG 2189
 QY 1408 cglttcccccgtgaagtcctccctgctgagcctctctgttccagaccctgacgtaccgat 1467
 Db 2190 CGTTTCCCTTGAAGAGTCCCTCGTGCGCTCTCTGTTCCGACCTGCGCCTTACCGGAT 2249
 QY 1468 acctgtccgcttctcccttccggaagcggtgagcttctcaatgctcaagctgtagat 1527
 Db 2250 ACCTGTCCGCTTCTCCCTTCCGGAAGCGTGGCGCTTCTCAATGCTCAGCTGTAAGT 2309
 QY 1528 atctcaagtcggtgtagtgcgtctcgaagcttgagctgtgtgcaagaaaccccgatc 1587
 Db 2310 ATCTCAATTCGCTGATAGTGTCTTCCCTCAAGCTGGGCTGTGTGCAGAAACCCCGTTT 2369
 QY 1588 agcccgacgcgtgcgcttaccgtaactatcgcttgaagtcacacccgtaagacag 1647
 Db 2370 AGCCCGACCGCTGCGCTTATCCGCTAATATGCTGTTGATGCTCAACCCGTAAGACAG 2429
 QY 1648 acttatcgcaactggcagcagcagcagctgttaacaggtatagcagaaggtatgtagcg 1707
 Db 2430 ACTTATTCGCCATGCGACGACGACCTGTAAAGGATTACAGGAGGATGTAAGCGC 2489
 QY 1708 gtgtcaagagttcttgaagttgttgactaacagcgtcactagaaagacagatattg 1767
 Db 2490 GTGCTACAGAGTCTTGAATGTGTGCTTACGCTACGCTACACTAGAGACAGTATTG 2549
 QY 1768 gtatctgcgtctgtcgaagccagttacacttcgaaaaagattgtgactcttgatccg 1827
 Db 2550 GTATTCGCGCTGCTGACGACCCAGTTACCTCGGAAAAAGATTGGTAGCTTGTATCCG 2609
 QY 1828 gcaaaaaaacacgcgtgtgagcggtgttlttcttgaagcagcagatatacagca 1887
 Db 2610 GCAAAACAAACCGCTGAGGAGGAGGTTTGTGTTGTTGCAAGCAGCAGATTAACCGCA 2669
 QY 1888 gaaaaaagatctcaagaagatccttgaatcttcttcttgaaggggtctgaagcctatgga 1947
 Db 2670 GAAAAAAGGATCTCAAGAAGATCTTGTATCTTTCTACGGGCTGACAGCTCAGTGGGA 2729
 QY 1948 acgaaactcaagtttaaggaatttggtcatagcaggtatacatattgaattataga 2007
 Db 2730 ACGAAAACTCACGTTAAGGATTTGTGTCATGAGATTATCAAAAAGATCTTCACCTAGA 2789
 RESULT 8
 LPLPREP
 LOCUS 2093 bp DNA linear BCT 24-APR-1996
 DEFINITION Plasmid pLp1 (from Lactobacillus plantarum CCM 1904) replication
 protein (rep) gene, complete cds.
 ACCESSION M31223
 VERSION M31223.1 GI:149684
 KEYWORDS replication protein.
 SOURCE Plasmid pLp1 (tissue library: CCM 1904) DNA.
 ORGANISM
 1 (bases 1 to 2093)
 REFERENCE
 AUTHORS Boula,A., Brinzel,F., Frey,L., Kammerer,B., Belarbi,A.,
 Guyonvarch,A. and Hubert,J.C.
 TITLE Structural organization of pLp1, a cryptic plasmid from
 Lactobacillus plantarum CCM 1904
 JOURNAL Plasmid 22 (3), 185-192 (1989)
 MEDLINE 90222315
 COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
 by B.Kammerer, 11-JAN-1990.
 FEATURES
 source 1..2093
 Location/Qualifiers

DEFINITION Plasmid pc3011 (from L.plantarum) ORF R, complete genome.
ACCESSION J03319
VERSION J03319.1 GI:144180
KEYWORDS complete genome.
SOURCE Plasmid pc3011 DNA.
ORGANISM Plasmid pc3011
plasmids.
REFERENCE 1 (bases 1 to 2140)
AUTHORS Skaugen,M.
TITLE The complete nucleotide sequence of a small cryptic plasmid from
Lactobacillus plantarum
JOURNAL Plasmid 22 (2), 175-179 (1989)
MEDLINE 90160873
COMMENT Draft entry and computer-readable sequence for [1] kindly submitted
by M.Skaugen, 19-NOV-1989.

FEATURES
source
1..2140
Location/Qualifiers
/organism="Plasmid pc3011"
/plasmid="plasmid pc3011"
/specific_host="Lactobacillus plantarum"
/db_xref="taxon:2535"
63..85
rep_origin
/note="plus-strand origin of replication"
145..150
-35_signal
167..172
-10_signal
361..366
RBS
/note="ribosome binding site"
372..1325
CDS
/note="ORF R"
/codon_start=1
/transl_table=11
/protein_id="AA91566.1"
/db_xref="GI:144181"
/translation="MSEIFEDKTEKNGKVPMPREKRIENRYAEYLAILEKPHADYRG
CGEVLEFRKIGELKLYQTFWFCRKRCPLCNMRKSKNSQKQIITAEVAREPKRF
LEFLTLVKNHSAEELKVSRLALTKAFNKLRYKTKNLGLRSTETIVNEDDSY
NOHLHLVHVSSEYFNSSNNYLAQAEMALMODALVDEPVPVHVAVANKRGIDS
LQASAEETAKYEKVSADYMTADDERNLVYIKNLEYALACTROISYGLGQIKODKL
EDVENGDVLHVGSEDYTKQEMAEAEVAVAKMDFKNQNYFM"
1337..1377
misc_signal
/note="transcription termination signal"
repeat_region
1469..1706
BASE COUNT 640 a 332 c 459 g 709 t
ORIGIN

Query Match 14.7% Score 1191.6; DB 1; Length 2140;
Best Local Similarity 94.5% Pred. No. 4.8e-272;
Matches 1279; Conservative 0; Mismatches 69; Indels 6; Gaps 4;

QY 2827 aaaaagcgtcagcggttttgcgtatcgtcttttggggttaaaaagc-caattttg 2885
|||||
Db 1 AAAAGCAGTTCACGCGTTTTCGCTGATCTGCTTTTGGGTTTAAACGCCAATTTTGG 60
QY 2886 ggcgtttctctatcttgaactatagaacaactagtttttaaaatcaagcttgat 2945
|||||
Db 61 GCGTTTCTTATCTTATCTGATATATAGCAACAAGTATTTTAAATCAAGCTTGAT 120
QY 2946 taagcttaattggcgtgtatcattatattataggcttttggtatattataggta 3005
|||||
Db 121 TAGCCTTGATTAGCCTTGTTCTTCCTGATTTTATAGCTTTCCGTTATTTAGCGTTA 180
QY 3006 taattggttgaagaagaacaataaacaacccagctgcaaatctcctagttggcgct 3065
|||||
Db 181 TAAATTGCTGAAGAAGACAAATTAACCAACCGTGAACATCTCCGATTGG--CGT 237
QY 3066 cggaaacagtgagttgattatcatcttgcgatt-tatagcctatctctaggggaaagccct 3124
|||||
Db 238 CCGAGCACGTGAGTGTATTCTTTTCATTGTGAGCCTTATTTAGTGTAAGAACCT 297
QY 3125 atgattgcaaggttatagcttatgt-aaaaagatagtagcctcctcacgttgataaac 3183
|||||

Db 298 TTGCTGCAAGGTTATATAGCTTATTAAGAAAAAGATAGTACGCTCTCTCACGTGTAAGC 357
QY 3184 tgaaggagctttttatgtcgaataatttgaagataaaactggaataatggcaagttagac 3243
|||||
Db 358 GGAAGAGGCTTTTATGTGCAAAATTTTGAAGATAAAAGTGAAGAAAGCAAGTTAGAC 417
QY 3244 ctggcgagaagcgaagaattgaaatgtgcgcattgcgcaataattggcaatctagaat 3303
|||||
Db 418 CTGGCGAGAACGGAAGATTGAAAATGCTGTTATGCCGAATTTTGGCAATCTTAGAAT 477
QY 3304 ttaacgycacatgatlgtacggygtgtgtgtgaagtttgcgttttcgtlaagattggcg 3363
|||||
Db 478 TTAACGGCCACATGATGTACGGGTTGTGCTGAAGTTTGGCTTTGCTGAAGATTGGTG 537
QY 3364 agcaactaaactttatcaacgtgtgttgcataaagattgttccattgtgtaatt 3423
|||||
Db 538 AGCAGTTAAACCTTTATCAACACGCTTTGTGATTAACCATTAATGTCATTTGTATAT 597
QY 3424 ggaagagcatgaaaaaactcgagccagttaaaacaattatttgcggaagcagttgcaa 3483
|||||
Db 598 GGAAGAAAGACATGAAGAAATTCAGCCCAATTAAACAAATTTATTCGGAGACAGTTGCA 657
QY 3484 gagaagcctaaagagcgttttgcgttttaacttaaccglttaaaacgctcattcagcag 3543
|||||
Db 658 GAGAGCCTAAAGACGCGTTTGTGTTTAACTTTAAACCGTTAAACGCTCATTTACAGCAG 717
QY 3544 aggaattaaaagtgcttttaagagcttgaacttaagccttaagcttaactcgcctata 3603
|||||
Db 718 AGGAGTTAAAGTGCTTTTAAGAGCTTAAAGCTTAAAGCTTAAATGCTAATGCTTATA 777
QY 3604 aaaaagtgactaaaatttatttggttatttaacttaacgtaacggaataccgttaatgaa 3663
|||||
Db 778 AAAAGGACATTAATAATTTATGCGTTATTTAGCTTCAACGGAATTAACGTTAATGAC 837
QY 3664 aagaaggtcatatcaatcaaacacttgcagatgtgtcgtgttgaataatcaagtattta 3723
|||||
Db 838 AAAGTGGCTATTTAAACACATTTTACATTTTGTGTTTGTGAAAGTAGTATTATTTA 897
QY 3724 agaattcaaatattatttgaacaaagcagaattgggcgaataattatgcaaaagccttga 3783
|||||
Db 898 AGAATTCAATTAATTATTATTTAGCACACAGCAATATGGCAAAATTTATGCAAAAGCTTTAA 957
QY 3784 aagttgattagagcctgtgtgtcatgtgcagctgttlaagcttaacaaacgtaagaa 3843
|||||
Db 958 AAGTTGATTTATGAGCCCTGTGTCATGTGCAGCTGTTAAAGCTAACCAACGCAAGAA 1017
QY 3844 ctgactctttgcgaagctagtgccggaagaacgycggaataatcaggtataatcgcgtatc 3903
|||||
Db 1018 CTGACTCTTTTGGCAAGCTAGTGCCGAAGAAAGCGCGAAATTCGAGGTAAATTCAGCTGATT 1077
QY 3904 atatgacgctgattgatagaacggttaatttgggtgtatgtaaaaatttgaatgacgtctag 3963
|||||
Db 1078 ATATGACGCGTATGATGACGCTAATTTGCTGCGATTAATAAATTTGCGATATGCCCTTAG 1137
QY 4024 ttgaagatgttgaagaatggtgattttagttcatgttggcgatgtaagattacccaagagc 4083
|||||
Db 1198 TTGAAGATGTTGAAGAAAGGATTTAGTTATGTTGCGCGATGAAGATTACCAAGAGAGC 1257
QY 4084 aatgtgaagctgcggaagaagctgtgcgcaaatgggaatttaataaacaattatttta 4143
|||||
Db 1258 AATGTGAAGCTGCGGAAGAGTGTGCGCAAAATTTGCGATTTTAATAACAAATTAATTTTA 1317
QY 4144 ttctgtaaaagagaatgtcaggaatagatcccg 4177
|||||
Db 1318 TTTGTTAAAGAGATGTACAGATATGATTAACCG 1351
RESULT 10
AF346623

LOCUS	AF346623	4699 bp	DNA	circular	SYN 30-AUG-2001
DEFINITION	RAGE vector PRIG20, complete sequence.				
ACCESSION	AF346623				
VERSION	AF346623.1	GI:15383986			
KEYWORDS					
SOURCE	RAGE vector PRIG20.				
ORGANISM	RAGE vector PRIG20.				
REFERENCE	artificial sequence; vectors.				
AUTHORS	1 (bases 1 to 4699)				
	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,				
	Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,				
	Whittington,J., Ierner,L., Costanzo,D., McElligott,K., Booser,S.,				
	Mays,R., Smith,E., Veloso,N., Kikla,A., Hess,J., Cotrehn,K., Lo,K.,				
	Offenbacher,J., Danzig,J., and Ducar,M.				
	Creation of genome-wide protein expression libraries using random				
	activation of gene expression				
	Nat. Biotechnol. 19 (5), 440-445 (2001)				
TITLE	212327151				
JOURNAL	11329013				
MEDLINE	2 (bases 1 to 4699)				
PUBMED	Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,				
REFERENCE	Cain,S., Dahl,T., Thornton,M., Ramachandran,R., Whittington,J.,				
AUTHORS	Ierner,L., Krashoc,D., McElligott,K., Clark,S., Mays,R., Smith,E.,				
	Veloso,N., Hess,J., Cotrehn,K., Lo,K., Offenbacher,J., Danzig,J.				
	and Ducar,M.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-FEB-2001) Athersys, Inc., 3201 Carnegie Ave.,				
	Cleveland, OH 44115, USA				
FEATURES	Location/Qualifiers				
source	1..4699				
	/organism="RAGE vector PRIG20"				
	/db_xref="taxon:161237"				
BASE COUNT	1062 a	1363 c	1275 g	998 t	1 others
ORIGIN					

Query Match	14.2%	Score 1152.4	DB 12	Length 4699
Best Local Similarity	76.4%	Pred. No. 1.le-262		
Matches 1524	Conservative	0	Mismatches 396	Indels 74
				Gaps 6
QY	23	gttcgacattgattatgactagtaattaagtaacaatcaacggggtcattagttcata	82	
Db	140	gttgcgacattgattatgactagtaattttatagtcaatttaccggggctcatttagttcata	199	
QY	83	goccalatalagagtlccgcgyltacaataacthaagtaaatgcccgcctgctgaccgc	142	
Db	200	gcccatatattgagagttccggcgttacattaaacttaacgtaattggccgcctgactgacggc	259	
QY	143	ccaaagacccccgcacctagtcagttcaataaagaagtaagttcccaatgaatgaacgcaatag	202	
Db	260	ccaaagacccccgcacctagtcagttcaataaagaagtaagttcccaatgaatgaacgcaatag	319	
QY	203	ggaacttccattgaagtcataatgggtgagactttaacgtaaacgtcccaacttggcagttac	262	
Db	320	ggactttccattgacgtcattatgggtgagactttaacgtaaacgtcccaacttggcagttac	379	
QY	263	atcaagtgatataatgccaagtaagcccccattgaagtcataatgacggttaattggccg	322	
Db	380	atcagaattatcatatgaccgaagtcacgccccctattgaagtcataatgacggttaattggccg	439	
QY	323	cctggacattatggccagtaacataagtaacttatatggactttccctactttggcagtaacatactag	382	
Db	440	cctggacattatggccagtaacataagtaacttatatggactttccctactttggcagtaacatactag	499	
QY	383	tattagatcatgacttaaccatgagtcagatgcggttttggcaatgaatacaatgaggtgagat	442	
Db	500	tattagatcatgacttaaccatgagtcagatgcggttttggcaatgaatacaatgaggtgagat	559	
QY	443	agcgatttgaactcaacggggaattccaagtctccaccaccattgaagtcacatggagatttgt	502	
Db	560	agcgatttgaactcaacggggaattccaagtctccaccaccattgaagtcacatggagatttgt	619	
QY	503	tttggcaccaaaatcaacgggactttccaanaatgttcgtaaacact-----ccgcccc	554	

Db	620	TTTGACACCAAAATCAACGGGACTTTCCAAATGTCTGTAACAACTGCCATGGCCCC	679
Oy	555	attagcgaatggcggttagcgctgtacggtggaggtctatataagcaagctctgt	614
Db	680	GTTGACGCAATGGGCGGTGAGCGGTGACGGTGGAGGCTCTATTATACAGAGCGCTT	739
Oy	615	gctactagagaaccactctctactggtctatcgaaatataagctactactagga	674
Db	740	AGTAAACCGTCAAGATCCGTAATTCGTAGACACCTACTGATTAAGCGCATAGAGCCTC	799
Oy	675	gaccaaagctgtgtaaccagctcgatccactagtaacgycgcgcagtgctgt-gaat	733
Db	800	TGCAAACTGTCTTAGTGACAACTATTCGATTTTCCACACTTATTAAGACCGCATGTATT	859
Oy	734	ctgcagatatacatcacacttgcgcgcgcgtcgagcatatcatatagggccctattcat	793
Db	860	GTCACACGCTCATGTCATGACGCTCCGGGAGACAGACAAAGCCGACCATGGCTCGAGTATA	919
Oy	794	agtgtaacctaaatgtctagagctcgtgtgtacagccctgcagcttgcctcttagttgcagc	853
Db	920	CGACTCACTATAGGGCCACAGGTGTAAGTACTCGCTACTTTAAGAGAG-----	965
Oy	854	catctgtgtttgcacctcccccgctgcctctcttgaaaccttgaagggtgcacccactg	913
Db	966	-----GCCATCTGTGGCCATTTAGCGTGGAGAAAGAAAGT	1000
Oy	914	tccttctctaataaatgaggaaatgtcatgcatctgtctgtagtgtatctatctac	973
Db	1001	TTAAGAGACCGAANCAAGCGCTCATGACCCGAAAGTGGCGAGGCCGATCTTCCCATCG	1060
Oy	974	tgggggggtggggtggggcagaagacgaaggggagatgtggaaagacaataagcaggtcat	1033
Db	1061	GTGATGTGCGGATATATAGGGCGCAGCAACGCCACCTGTGGCCCGGTATCCGGCCACG	1120
Oy	1034	ctggagatgctgggtggtctatgtctctcgtggcggaagaaaccgcgtcgaatgaat	1093
Db	1121	ATGGG-----TCCGGCTTAGAGATCCACAGCGGGGTGTGTCGCATGA	1166
Oy	1094	cgcgcaacgcgcggggagagcggttgcgtatgtggcgctcttcgcgttcctcgtcac	1153
Db	1167	TGCGGTATGTGATATGTGCTCCAAATAGGGAAGGACGACGACTGGCGGGGCCAAAGC	1226
Oy	1154	tgcctgcctgcgtctggttgcgtctgcgtgcgcgcgcgtatagctactcaaaagcgt	1213
Db	1227	GGTGGACAGTGGCTCCG-ACAAGGGGTGGCGCATATGAATTTGATCAACGATATAGCGCT	1285
Oy	1214	aatgcgttatccacaagaatcagggataacgcagaagaacatgttagaanaagcgca	1273
Db	1286	AGATCTCTTGCTAAGTGTGAAATCTGTGCGAGC-----CATGTAGCAAAAAGGCA	1334
Oy	1274	gcaaaagccagaagaccgtataaaagccgcgctgtctgtgcgttttccataagctcgcgc	1333
Db	1335	GCAAAAGCGCAGAAACCGCTAAAAAGCCGCGTGTGCGGCTTTTCCATAGGCTCCGGCC	1394
Oy	1334	ccctgcagcagcatacaanaaatcgaagctccaagctagaggtgtgcgaaccgcgaagact	1393
Db	1395	CCCTGACGAGCATCAAAAATGACAGCTTCMACTCAGAGGTGGCAAAACCGACAGACT	1454
Oy	1394	ataaagataccagcgctttcccccctggaaagctccctcgtgcctctcgtltcogaacct	1453
Db	1455	ATAAAGATTACAGCGCTTTTCCCTCTGGAAGCTCCCTCGGCGCTCTCTGTTCCGACCT	1514
Oy	1454	gcgcgttaacggataactgtgcgcgttttcccttccttcggaaagcgtgggcctttcgaatg	1513
Db	1515	GCCCGTTAACCGAATCCTGTGCCCTTTCTCCCTTGCGGAACGAGGCGCTTCTCATAG	1574
Oy	1514	ctcagcgttagatcatctcaagttcgggttaggtgcgttcgcctccaagcctgggcgtgtgca	1573
Db	1575	CTCACGCTGTAAGTATCTCAGTTCGGGTGAAGTCTGTCCCTCCAAGCTGGGCTGTGTCA	1634
Oy	1574	cgaacccccggttaagcccgacccgctgcgccttatccgtlaactatcgttcttagtcaaa	1633


```

RESULT 12
PCDNA32EO 3986 bp DNA linear SYN 16-AUG-1995
LOCUS      Cloning vector pcdna32EO DNA.
DEFINITION X00639.1 GI:949972
ACCESSION  X00639
VERSION     cloning vector; expression vector; multiple cloning site; plasmid.
KEYWORDS    synthetic construct.
SOURCE      ORGANISM
REFERENCE   1 (bases 1 to 3986)
AUTHORS    Peters,H., Hundhausen,T., Kroenke,M. and Margel,M.
TITLE      A new small sized high-level eukaryotic expression vector
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 3986)
AUTHORS    Peters,H.
TITLE      Direct Submission
JOURNAL    Submitted (07-AUG-1995) H. Peters, Inst. f. Immunologie,
           Michaelisstr.5, D-24105 Kiel, FRG.
           Related sequences: M21295 and K03104.
COMMENT     Location/Qualifiers
FEATURES    source
            1..3986
            /organism="synthetic construct"
            /plasmid="pcDNA32EO"
            /db_xref="taxon:32630"
            1..2125
            /note="cloning vector (pcDNA3) (Invitrogen)"
            misc_feature
            889..994
            /note="multiple cloning site (MCS)"
            misc_feature
            2126..2796
            /note="cloning vector (PzeosV) (Invitrogen)"
            misc_feature
            2797..3986
            /note="cloning vector (pcDNA3)"
BASE COUNT 900 a 1057 c 1054 g 975 t
ORIGIN
Query Match 13.3%; Score 1082; DB 12; Length 3986;
Best Local Similarity 100.0%; Pred. No. 6e-246;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gatatacggcagatatacagcgttgcattgattatgactagttatataatgaatca 60
Db 210 GATGACGGCGGACATATACCGCGTTGACATGATTATGACTAGTTATTAATGAATCA 269
QY 61 attacggggtcattagttacatagccatataatgagttccggttacataactacgta 120
Db 270 ATTACGGGTCATTAGTTCATAGCCCATATATGAGTTCGCGTTACTATTAATGAGSTA 329
QY 121 aatggccgcctggtgacggcccaagaccccgcccatgagtgacgtacataatgacgtat 180
Db 330 AATGGCCCGCTGCTGACCGCCCAAGACCCCGCCCATGACGTCATATATGACGTAT 389
QY 181 gtcccatagtaacgcaatagagacttccattgaacgtcaatggtgacgtatgaacg 240
Db 390 GTTCCCATAGTAACGCCCAATAGGAGCTTTCATTGACCTCAATGGGTGACTATTTAGG 449
QY 241 taacatcgccacttggagcatcatcaatggtatcatatgacgaagtaacgccccattgac 300
Db 450 TAAACTGCGCCACTTGGCGAGTACATCAAGTGTATCATATGCAAGTACGCCCTATTTAC 509
QY 301 gtcaatagagtaaatgacggcgtgcatatgtcccaagatacagacttatggacttt 360
Db 510 GTCAATAGCGGTAATAGCCCGCGCTGCGATTATGCGCGTATGAGACTTTT 569
QY 361 cctacttggcagcatcatcattgattgacatcgctatcatcagtgatgacgttttgg 420
Db 570 CCTACTTGGCGAGTACATTAGCTATTCGCTATTCATCAATGATGATCGGCTTTTGG 629
QY 421 cagtaacataatggcggtggaatagcggttgaactcaacggggaattccaagctccacccc 480
Db 630 CAGTACATCAATGGGCGGTGATAGCGGTTTGACTCAGCGGAGATTCCCAAGTCTCACCCC 689

```

```

QY 481 attgacgtcaatgggagattgttttggcaccacaatacaacgggactttccaanaatgct 540
Db 690 ATTGACGTCATATGGGAGATTGTTTGGCACCAAAATACAGCGGACTTCCAAAATGTCGT 749
QY 541 aacaattccgcccattgagcgaatggcggtgagtggttgaagttgggggtttatata 600
Db 750 AACAACTCCGCCCATTTGACCAAAATGGCGGTAGGGGTGATCGGTGGAGTCTATATA 809
QY 601 agcagagctctctgtgccaactagagaacccaactgcttactgtgcttaccgaataatgac 660
Db 810 AGCAGAGCTCTCTGTGCTACTAGAGAACCCACTCTTACTAGCGTTATCGAAATTAATACG 869
QY 661 actcaataatggagaccacaagcttggtaaccagctcgatccatcagtaagtaacggcgca 720
Db 870 ACTCACTATAGGAGACCCCAAGCTTGGTACCGACTCGATCCACTGATTAACGCCGCCCA 929
QY 721 gtgtcgtggaatttgcagatattccatacaactgctggccgctcgagcatgacatcagag 780
Db 930 GTGTGCTGGAATTTGTGACAGATATCATACACTGCGCGCGCTGACATGATCTAGAG 989
QY 781 ggcctattctatagttcaccctaataatgtagagctgctgatalcagctcgaactgtgct 840
Db 990 GGCCCTATTTCTATAGTGTACACTTAATGCTAGACTCGTGTGATGACCTGAGCTGCTCT 1049
QY 841 tctagttgcagccatctgtgtgttgcctccctcccgctgcttctccttgacccttgaagt 900
Db 1050 TCTAGTTGCCAGCCATCTGTTGTGGCCCTCCCGCTGCTCTTACCCCTGGAAGT 1109
QY 901 gccactcccaactgtccttccctaataatgagaatgcatgcatgctgtctgtagtag 960
Db 1110 GCCACTCCCACTGTCTTCTCTATTAATGAGGAATTCGATGCGATTTGCTGATAGAG 1169
QY 961 tgtcattctattctgggggttgggttggcgaggacagcaaggggggagattgggaagac 1020
Db 1170 TGTCACTTCTATTTCTGGGGGCTGGGGTGGCGGACAGACAGCAAGCGGAGATTTGGGAAGAC 1229
QY 1021 aatagcagcagctgtctgggagtggtggtggtctgtggtcttggcggaagaacagc 1080
Db 1230 AATAGCAGCGCATGCTGGGGATGCGGTGGCTCTATGCGCTTGAGCGGGAAGAACAGC 1289
QY 1081 tg 1082
Db 1290 TG 1291

```

```

RESULT 13
AX319694 5446 bp DNA linear PAT 14-DEC-2001
LOCUS      AX319694
DEFINITION Sequence 5 from Patent WO0181614.
ACCESSION  AX319694
VERSION     AX319694.1 GI:17901350
KEYWORDS    synthetic construct.
SOURCE      ORGANISM
            synthetic construct.
            artificial sequence.
REFERENCE   1 (sites)
AUTHORS    Leng,J.
TITLE      Cell proliferation assay
JOURNAL    Patent: WO 0181614-A 5 01-NOV-2001;
            Chemicon International (US)
FEATURES    location/Qualifiers
            1..5446
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="pcDNA3 vector sequence"
BASE COUNT 1255 a 1417 c 1390 g 1384 t
ORIGIN
Query Match 13.3%; Score 1082; DB 6; Length 5446;
Best Local Similarity 100.0%; Pred. No. 6.2e-246;
Matches 1082; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY	1	gatytaaggagccaaataaagcgtttacgtttacgtttatbtacacgtattatataaagaacca	60
Db	210	GATGTAGCGGCCAGATATTACGGCTTGACATTGATATATTACATAGTATTATTAATAGTAATCA	269
OY	61	attacgggtcatatagttcataagcccatatattagttccggtttacataacttaccgla	120
Db	270	ATTACGGGGGTCATTAGTTTCATTACCCCATPATATAGGATTCGCGGTTATCATTAACCTTAGGTA	329
OY	121	aatygccgccttgcttgaaccgcccacagaccccgcccatgtacgtlcaataatgaaglat	180
Db	330	AATGGCCCCGTGGCTGACCCGCCAACAGACCCCCGCCCATTTACAGTCATAATATGACGAT	389
OY	181	gttcccaatgtaaagcccaataggagactttccatattgacgttcaatgggttgacatttaag	240
Db	390	GTTCCCATAGTAACGCCAATATAGGACTTTTCATTATGACGTCAATGGGTGACATTATTACGG	449
OY	241	taaacgtcccaacttggcagttacatacaagtgtalcatatgtccaagtaagcccccattac	300
Db	450	TAAACTGCCACTTGCGGACATCAAGTGATCATATGCAAGTACGACCCCTATTGAC	509
OY	301	gtcaatgaaggttaatbgycccgccctggcatltagccagttacatpacttatbgygacttt	360
Db	510	GTCATATAGCGTAAATGGCCCCGCTGGCATTTAGTCCCATACATGACCTTATGGGACATTT	569
OY	361	ccctacttggagttacatctacgtattagttacatgcgtattacatgtgtatgtggtttgg	420
Db	570	CTTACTTTGGCAGTACATCTAGTATTAGTCATGCTATTACATGCTATGCGGTTTGG	629
OY	421	cagttacatcaatgggctgtgataagcgggtttbgtactcaaggggatttccaaagttccaccc	480
Db	630	CAGTACATCATATGGCGCTGGATATGCGGTTTGATCACGGGGATTTCCAAAGTCTCCACCCC	689
OY	481	attgaactgaatggagtttggttttggacccaaataccaacggagactttccaaatgctgt	540
Db	690	ATTGACTCATATGGGAATTGTTTGTGGCACCAAAATCAACGGGACTTTCCAAATATGCTGT	749
OY	541	aacaactcggccccaattgacgcgaatbgygcgtatgacggttgaagctgtatcgtttatata	600
Db	750	AACAACTCCCCCATGTGACGAAATGGCGGATGAGCGTGTACGGTGGGAGGTCTATATTA	809
OY	601	agcgaagacttcttggcgttaacttagaagaaccacgtcttaagctttatcgaaattraatcg	660
Db	810	AGCAGAACTCTTGCGCTTAAGTATAGAAACCCACTGCTTACTGGCTTATGCATAATTAATACG	869
OY	661	acttaccataagggagaccaccaagctttgtltagccgagctcggatccacttagtlaacggccga	720
Db	870	ACTCACTATATAGGAGAACCCAAAGCTTGTAOCGAGTCGGATCCACTAGTAACGGCGCCA	929
OY	721	gttggctcgggaattctctgacagataltcaatacacttgcgcgtctgaagcatgcatctagag	780
Db	930	GTTGGCTGGGAATTTCTGCAGATATTCATCAACACTGGGGCGCGTCCAGAGATGATCTATAG	989
OY	781	ggcccattctatatagtttcaaccfaaatgtctagagctcgtgtatcaagccctgaacttgcct	840
Db	990	GGCCCAATTCTTATAGTGTACACTTAATGCTAAGAGCTCGCTATACAGCTCGACTGTGGCT	1049
OY	841	tctagttgcagccatctgttgttcttgcctcccccgcgttgccttccttgaaccttgaaggt	900
Db	1050	TCTATGTTGCACAGCATCTGTATGTTTCCCTCCCTCCCGTCCCTTCAACCTCGGAAGGT	1109
OY	901	ggccactcccaacttcccttcccttaataaattgggaattgatactgcatgttctgaatagg	960
Db	1110	GGCACTCCCACTGCTCTTTCTTATTAATAATGAGGAATATGCATCCATATGCTGTGATGG	1169
OY	961	tgatcatctatctctbgyggggttgggttgcgggacgaagaagaagggggagatattggaagaac	1020
Db	1170	TGTATATTCTATTTCTGGGGGCGTGGGGGAGAGACAGCAAGGGGAGGATTTGGGAAGAC	1229
OY	1021	aataggagagatctcggggatgtggtgtggtgtctatagtgcttctgaagcgggaagaaccgcg	1080
Db	1230	AATATGCAAGGATCCTGGGGATGCGTGGGCTCTATGCGCTTCTGAAGCGGAAAGAACCCAGC	1289
OY	1081	tg 1082	

Db	1290	TG	1291	11
RESULT	14			
LOCUS	AR098191	4026 bp	DNA	linear
DEFINITION	Sequence 19 from patent US 6074850.			PAT 14-FEB-2001
ACCESSION	AR098191			
VERSION	AR098191.1	GI:12807448		
KEYWORDS				
SOURCE	Unknown.			
ORGANISM	Unknown.			
REFERENCE	1 (bases 1 to 4026)			
AUTHORS	Antelman,D., Gregory,R.J. and Wills,K.N.			
TITLE	Retinoblastoma fusion polypeptides			
JOURNAL	Patent: US 6074850-A 19 13-JUN-2000;			
FEATURES	Location/Qualifiers			
source	1..4026			
BASE COUNT	978 a 1020 c 983 g 1045 t			
ORIGIN				
Query Match	13.2%	Score 1068.6;	DB 6;	Length 4026;
Best Local Similarity	69.7%	Pred. No. 9.1e-243;		
Matches 1918:	Conservative 0;	Mismatches 89;	Indels 746;	Gaps 5;
QY	1	gagtcggcgccagataaacgcttgacattgatttgacattgataatgaatca	60	
DB	210	GATGTACGGGGCCAGATATACGCCGTTGACATTGATTTATGACATTATTAATGATATCA	269	
QY	61	attcgagggtacatgaatcctatagcccatatagatgcggttccattcaacttcgcta	120	
DB	270	ATTACGGGGGTATTATGTTTCATATACGCCCATATATGAGTTCGCGCTTACATTAATTAATGATCA	329	
QY	121	aatggccgccttgctcgtacacgcccacagaccccgcccatctgacgtcgtcaataatgacgtat	180	
DB	330	AATGGCCCGCGCTGCTGACCGCCCAACGACCCCGCCCATTTGACGTCAATTAATGACGTAT	389	
QY	181	gtcccatagttaagcccaataaggactttccattgacgtcaatgggtggactattacgg	240	
DB	390	GTTCCCATATGTAATACGCCAATAGGAGACTTTCATTTGACGTCAAGGGGTGACATATTTACGG	449	
QY	241	taaacctcccacattggagcagfacatcaagtgtatcatatgccaagtacgccccattgac	300	
DB	450	TAAACTGCCCCATTGGGAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTATGAC	509	
QY	301	gtcaatgacggttaaatgagccgccttgccattatgcccagtlacatgacattatggagactt	360	
DB	510	GTCATATGACGGTAAATGCGCCCGCTGCGCATTTATGCCACATGACATTAATGCGACTTT	569	
QY	361	cctacttggagatgacattcctagctattagatcatgcgtcattaccatggtagtgggttttgg	420	
DB	570	CCCTACTGGGAGTACATCTACGTATTAAGTATGATGCTATTAACCATGGGTGATGGGTTTTGG	629	
QY	421	cagtlacatcaatggcgctgtgataagcggtttgaactacacgggagatttccaagctccaaccc	480	
DB	630	CAGTACATCAATGGGCGGTGATATGCGGTTTGACTACGGGGAATTTCCAAAGTCTCCACCCC	689	
QY	481	attgacgtcaatgggagattgttttggcaccaaatccaacggagactttccaaaatgctgt	540	
DB	690	ATTATGCTCATATGGGAGCTTTGTTTGGCACAAATTAACAGGGGACTTTCCAAATAGTCTT	749	
QY	541	aacaacccgcccccatgacgcaaatggcggttaggcgtgtacacggtggaggttcataata	600	
DB	750	AACAACCTCCGCCCATTTGACGCAATAGGGGGTATGGCGTATACGGGTGATATTA	809	
QY	601	agcagagctctctggttaactagagaaccactgtctacttggcttatcgaaatataacg	660	
DB	810	AGCAGACCTCTTGCGCTTAATAGAGAACCCACACGCTTACTTGCTTATCAAAATTAATACG	869	

[illegible]

Db	1950	CTGGGGTGCCTATATAGTAGCTAACCTACATTAATTGGCGTTGGCCTACTACGCCGCTTT					2009
QY	1063	-----gagcggaagaacacagctgcatatgaaatcgcccaacgcgcggaagagag					1114
Db	2010	CCAGTCGGGAAACCTGTGTGTGCACACTGCATTAAATGAATMGCCCAACGCGGGGAGAG					2069
QY	1115	cggtttgcgtatgtaggcgtcttcgcgtcttcgtctcaatgactgcgtgcgtgcgt					1174
Db	2070	CGGTTTGGGTATTTGGGGCCCTCTTCCGCTTCCCGCTCACTGCATCGTCGCGCTGCTGT					2129
QY	1175	tgggtctggcggaaggggtatcagctcactcaaaagcgataagggattatccaaagatc					1234
Db	2130	TGCGCTGGCGGCGAGCGGATACGCTCACTCAAAAGCGGTAATACGGTATTCACAGAAATC					2189
QY	1235	aggggataacgcgaagaaagacatgtgacaaagggcagcaaaagggccaggaacgtaa					1294
Db	2190	AGGGATTAACGACGAGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCTTAA					2249
QY	1295	aaaggccgcgtgtgcgtggcggttttccataagctccgccccctgacggcatcaaaaaa					1354
Db	2250	AAAGGCGCGGTGTGCGGGGCTTTTCCATAAGGCTCCGCCCTCTGACGAGATCAAAAAA					2309
QY	1355	tgaagcgtcaagtcagaggttgcagaaacccgaacagactataaataaccagcgcttcc					1414
Db	2310	TGAGAGCTTCAAGTCAGAGGTGGCGAAACCCGACAGAGACTATTAAGATACAGCGCTTCC					2369
QY	1415	cccttgaagctccctcgctgagctctcctggttccgaacctgcccgttaccggaatacgttc					1474
Db	2370	CCCTGGAAAGCTCCCTCGTGGCGCTCTCGTTCGACCCCTGCGGCTTACCGGATACCTGTC					2429
QY	1475	cgcccttctcccttcgggaaagcgtgtggcgctttctcaatgctcaagctgttaagtactaag					1534
Db	2430	CGCCCTTTCCTCCCTTGGGGAAGCGTGGCCCTTCTCAATGCTCACGCTGTAGTATCTGAG					2489
QY	1535	ttcggttgaagtcggttcgctccaaagctgggcgtgtgtgacgaaccccccggttcaagccga					1594
Db	2490	TTGCGGTGAAGTCGTTCCTCCAACTGGGCTGTGTGTGACAGAACCCCGCTTACGCCGA					2549
QY	1595	ccgctgcgcctatccogttaaactatcgcttcttgatgctcaaccgcgtaagaacacgactatc					1654
Db	2550	CCGCTGCGCCTTATCCGGGAACCTATGCTGTGATGCCAACCCGGTAAAGACAGACTTATTC					2609
QY	1655	gccacttgcagcagccactctgttaacaagattagcagaagcgtgtatctagcgcgtgtctac					1714
Db	2610	GCCACTGGGAGGACGACCTGCTGTAACAGATTATGCAAGACGAGATGTGTAAGGGGGCTTAC					2669
QY	1715	agaagttctgaagtggtgacctaaactaagcgtcaacactagaagaagaagatattgtrattg					1774
Db	2670	AAGAGTCTTGAAGTGTGTGCTTAACCTACGGGCTTACCTGTGAAAGAGATATTTGGTATCTG					2729
QY	1775	gcgcttgcgtgaagccagttacaccttcggaaaaaagattgtagctcttgatcccggaanaa					1834
Db	2730	CGCTGTGCTGAAGGACGAGTACTCTTGGAAGAAAGATTTGGTATCTTGTATCCGGCAAAACA					2789
QY	1835	aaaccaacgcgtgtagcgtgtgttcttctgttgcagaacagcagatlcgcgcagaanaa					1894
Db	2790	AACCAACCGCTGCTAGCGGGGCTTTTGTGTTTGTGCAACGACGAGATTACGCGCAGCAAAAAA					2849
QY	1895	aggaatctcaagaagaatcctttgatcttcttctacggggctctgaagctcaagtggaaagaaa					1954
Db	2850	AGGATCTCAAGAAAGATCCCTTGATCTTTTCTTCAACGGGGCTGTGACGCTGACGTGAAGGAA					2909
QY	1955	ctcaaggttaaggaatttgcatacgaagcagataattgaattattaga 2007					
Db	2910	CTACGCTTAAGGAGATTGTGTCATGAGATTATCAAAAAAGATCTTACACTAGA 2962					
RESULT 15							
LOCUS	BD009730	4026 bp	DNA	linear	PAT 31-JAN-2002		
DEFINITION	Tissue specific expression of retinoblastoma protein.						
ACCESSION	BD009730						

```

VERSION      BD009730.1 GI:18638103
KEYWORDS     JP 2001503638-A/4.
SOURCE       unidentified.
ORGANISM     unidentified.
REFERENCE    1 (bases 1 to 4026)
AUTHORS      Antelman,D., Gregory,R.J. and Wills,K.N.
TITLE        Tissue specific expression of retinoblastoma protein
JOURNAL      Patent: JP 2001503638-A 4 21-MAR-2001;
              CANOT INC

COMMENT
OS   Unidentified
PN   JP 2001503638-A/4
PD   21-MAR-2001
PR   13-NOV-1997 JP 1998522958
PR   15-NOV-1996 US 08/751517,14-FEB-1997 US 08/801092 P1
CC   DOUGLAS ANTELMAN,RICHARD J GREGORY,KENNETH N WILLS PC
CC   C07H21/04,C07K5/00,A61K38/00,A61K35/12
CC   Strandedness: Single;
CC   Topology: Linear;
FH   Key
FT   source      1..4026
                   Location/Qualifiers
                   1..4026
                   /organism="unidentified"
                   /db_xref="taxon:32644"

BASE COUNT   978 a 1021 c 982 g 1045 t

ORIGIN
Query Match      13.1%; Score 1067; DB 6; Length 4026;
Best Local Similarity 69.6%; Pred. No. 2,2e-242;
Matches 1917; Conservative 0; Mismatches 90; Indels 746; Gaps 5;

QY 1 gatgtacgggcccagataacagcggttgcacattgattatgactagttatataatgaatca 60
   |||
DB 210 GATGTACGGGCCAGATATAGCGCTTGACATGATTATGACTGATTATATATGATATCA 269
QY 61 attacggggcattagttcatagcaccatataatagagttccgggttaccataactacgta 120
   |||
DB 270 ATTACGGGGCATTAAGTTCTATAGCCCATATATGAGAGTTCGGCTTACATATACGGTA 329
QY 121 aatggccgccttgctgacccgccaacgaccccccgcacattgacgtaacaaatgacgtat 180
   |||
DB 330 AATGGCCCGCTGCTGACCGCCCAACGACCCCGCCCATTTGACGTCATATATGACGTAT 389
QY 181 gtcccatagtaacgcaataagagcttccattgacgtcaatgaggtgtgacattacgg 240
   |||
DB 390 GTTCCATATGTAAGCCAAATAGGAGCTTTCATTTGACGTCAATGGGTGACTTTTAA 449
QY 241 taaactgcccacttggaagtaacaaatgtaacatataatgcaagtaacgccccctattgac 300
   |||
DB 450 TAAACTGCCACTTGGCAGTACATCAAGTATCATATGCAAGTACGCCCCCTATTGAC 509
QY 301 gtcaatgacgtaaatgcccgcctggaattatgcccagtaacatgacctatggaactt 360
   |||
DB 510 GTCAATGACGGTAATAGCGCCCGCTGCAATTATGCCAGTACATGACCTTATGGACTTT 569
QY 361 cccaactggcagtaactaagtaagtaagtaagtaagtaagtaagtaagtaagtaagta 420
   |||
DB 570 CCTACTTGGAGTACATCTACGATTATGATGCTATTATACATGATGATGATGATGATG 629
QY 421 caatatacatatgagcgtgagatagcggttgaactcaacggaatttccaagtcacaccc 480
   |||
DB 630 CATATATCATATGGGGGTGATAGCGGTTTGAATCAACGGGAGATTTCCAAGTCTCCACCC 689
QY 481 attgacgtcaatgagaggtttgttttggcaacaaatcaacggaacttccaacaaatgct 540
   |||
DB 690 ATTGAGGTCAATGGAGTGTGTTTGGCAACAAATCAACGGGAGCTTCCAAATATGCGT 749
QY 541 aaaaactccgcccattgacgcaaatgagggcggtgtaagcggtgagagaggtctatata 600
   |||
DB 750 AACAACTCCGCCCATTTTGAACGAATATGGCGGTAGGCGGTACGGTGGGAGGTCTATATA 809

```

```

QY 601 agcagaagctctcgtgcttaactagaagaccacactgcttactggtctatgaaatataacg 660
   |||
DB 810 AGCAGAGCTCTCTGCTTAACATAGAGAACCCACACTGCTTACTGCTTATCGAAATTAATACG 869
QY 661 acccactatagagagaccacgaacttg ----- 687
   |||
DB 870 ACTCACTATATAGGAGAGACCAAGCTTGGCGGGGTACCACCTCTCTCCGATCCGCTGTG 929
QY 688 ----- 687
DB 930 CGAGGCGCAGCTGTGGCTCGCGGTTAGAGAGAACTCTTCGCGGCTTTCCAGTACTC 989
QY 688 ----- 687
DB 990 TTGATCGGAACCCGCTGGCTCCGACGAGTACTCCGCCACCGAGGACCTGAGGAGT 1049
QY 688 ----- 687
DB 1050 CCGCATGACCGGATCGAAGAAACCTCTCGAGAACTGAAAAACAGAAATTATCTGTGA 1109
QY 688 -----tacgaagctcgatccactagtaacgagccgacgtg 723
DB 1110 AGTTAGTCTTTTGTCTTTTATTTTACAGTCCCGGATCCGGTGTGTGCAATCAAG 1169
QY 724 tgctggaattcgcagatat----- 743
DB 1170 AACTGCTCTCACTGATGATGTTGCTTACTTCTAGGCTGTAGCGAAGTTACTTCTGC 1229
QY 744 -----ccatcacactgcgcgc-----tcgaacatgcat 774
DB 1230 TCTAAAGCTCGCGAATTGATACCGCGCGCTGACGCTAGACGAATTCGGCTAGATA 1289
QY 775 ctgagagggcctatctatagtgtaactaaatgctcaagctcgcgtgacacgctgact 834
   |||
DB 1290 TCGATGGCCCAATTTCTATAGTGCACCTAAAGCTAGAGCTCGCATCAGCTCGACT 1349
QY 835 gtgcctcagttgacgacatctgtgtttgtccctcccgctgctccttgacccgt 894
   |||
DB 1350 GTGCCTTCTAGTTGCCACGACATCTGTTTGGCCCTCCCGCTTCCTTGACCTG 1409
QY 895 gaagtgcaactccacgtgcttccctaaataatgagaaatgcatgcattgctg 954
   |||
DB 1410 GAAGGTGCCACTGCCACTGCTCTTCTTAATAAAGAGAAATTCATCGCATGCTGCTG 1469
QY 955 agtaagtgcatctatctctgtggggtggtggtggtggtggtggtggtggtggtggtg 1014
   |||
DB 1470 AGTAGGTGTCATTCTATCTGgggggtggtggtggtggtggtggtggtggtggtggtg 1529
QY 1015 gaagacaatagcagcatgctggtga ----- 1040
   |||
DB 1530 GAAGACAATATGCCGAATATGACGACCAAGCGCCCAACCTGCCATCAGAGATTGCA 1589
QY 1041 ----- 1040
DB 1590 TTCCACGCCGCTTCTATGAAGGTTGGGCTTCGGAATCGTTTCCGGAGCGCGGCTG 1649
QY 1041 ----- 1040
DB 1650 GATGATCTCCACGCGCGGGAATCTATGCTGAGATTCTTGCCGCCACCCAACTGTTTAT 1709
QY 1041 ----- 1040
DB 1710 TCGAGCTTATATGTTACAAATTAAGCATATGACATCAAAATTTCCAAATTAAGCATT 1769
QY 1041 ----- 1040
DB 1770 TTTTTCATGCTATCTAGTTGTGTTTGTCAAACTCATATGATTTATCATGCTGTG 1829
QY 1041 ----- 1040
DB 1830 TATACCGTCCGACCTCTAGTAGCTTGGCGTAATCATGCTATGATGCTGTCTGTGTG 1889

```

Search completed: May 24, 2002, 06:25:19
Job time: 20231 sec

```
QY 1041 ----- 1040
Db 1890 AAATGTTATCCGCTCACAATTCACACACACATACAGACCCGGAACATAAAGTAAAGC 1949
QY 1041 ----- 1062
Db 1950 CTGGGGTGCCCTAATAGTAGTACCTACATTAATTCGCTTGCGCTCCTACTGCCCCGCTTT 2009
QY 1063 ----- 1114
Db 2010 CCACTCGGGAAACCTGTCGTCCAGCTGCATTAATGATCGCCACACGGCGGAGAGAG 2069
QY 1115 cggcttcgcatatggcgctctccgctccctcgctcactgactcgctcgctcgct 1174
Db 2070 CGGTTTCGATATGGGCGCTTCCTCGCTCCTCAGTACCTGCTGCGCTGCGT 2129
QY 1175 tcgagctcggagagcggtatcagctcactcaagaagcggtatcaggtatccagatc 1234
Db 2130 TCGGCTGGGCGAGCGGATATCCTCACTCAAGCGGATATACGTTATCCACAGATC 2189
QY 1235 aggggataacgcagaaagaacatgtagcaaaagccagcaaaagccagaaacgta 1294
Db 2190 AGGGGATTAACGCAGAAAGAAACATGTAGCAAAAGCCAGCAAAAGCCAGAAACGTTA 2249
QY 1295 aaagcgcgctgctgctgcttccatagctccgccccctgacgagcatcaaaaa 1354
Db 2250 AAAGCGCGCTTGCTGGGCTTTTCATAGGCTCCGCCCTCAGACGATCACAAAAA 2309
QY 1355 tcgagctcagatcagaaggtggcgaaacccgacaggaactataagataccagcgcttcc 1414
Db 2310 TCGAGGCTCAAGTACAGAGTGGCAAAACCCAGACAGACTATTAAGATACAGGCGTTCC 2369
QY 1415 cccctggaagctccctcgtcgtctcctgctcgaacctgcccgttacccgatacctgctc 1474
Db 2370 CCGTGAAGGCTCCCTCGTGGCTCTCGTCCGATCCGCTACCGGATACCTGTC 2429
QY 1475 cgccttcctccctcgggaagcgctggcgcttccatagctcagcgctgtagtctcag 1534
Db 2430 CGCCTTTCCTCCCTCGGAAACGCGGCTTTCATAGCTCACGCTGTAGTATCTCAG 2489
QY 1535 ttgggttaggtcgttcgctcgaagcgtggcgtgtgtgcaagaaaccccgctcagccga 1594
Db 2490 TTGCGGTAGGCTTTCGCTCAAGCTGGGCTGTGTGTCACGAAACCCCGTTCAGCCGA 2549
QY 1595 ccgctcgcccttacccgtaactactcgtctgtagtccaaacccggtlaagacagactatc 1654
Db 2550 CCGCTGCGCCTTATCCGCTATCGTCTGAGTCCAAACCCGGTAAGACACGACTATC 2609
QY 1655 gccactggcagcaaccactggttaacagatcagaagcgaggtatgtagcggtgctac 1714
Db 2610 GCCACTGGCAGCAGCCACTGGTAAACAGATTAAGCAGACCGAGTATGAGGGGTGCTAC 2669
QY 1715 agagtctcgaaggtgtagcctaactacggtacactagaaggaagcaatattggtatctg 1774
Db 2670 AGAGTTCTGTGAAGTGGTGGCTTAACGTACGCTACACTAGAAGGACAGATTTGGTATCTG 2729
QY 1775 cgcctcgtcgaagcaggttacctcggaaaaagagttgtagctctgaltccggcaaca 1834
Db 2730 CGCTCTGCTGAAGCCAGTTACCTTCGAAAAAGATTGGTAGCTTGATCCGCCAAACA 2789
QY 1835 aaccacgcgtgtagcggtggtttttgttgcagaagcagagattacgcgcaaaaaaa 1894
Db 2790 AACCAACGCGTGTAGCGGTGTTTTTTGTTTGAAGCAGCAGATTAACGCCAGAAAAA 2849
QY 1895 aggattcacaagaagatcccttgatcttctcaggggtcctacgctcagltgaacgaaaa 1954
Db 2850 AGGATTCACAAAGATCTTGTGATCTTTTCACGGGCTGTGACCGTCACTGGAAGAAAA 2909
QY 1955 ctcaagttaaaggaatttgatcagcgagatcacatatgtaattgatttaga 2007
Db 2910 CTCACGTTAAAGGATTTTGGTCATGAGATATCAAAAAAGCATCTTCACTCAGTA 2962
```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: May 24, 2002, 00:00:03 ; Search time 755.38 Seconds
(without adjustments)

18444.704 Million cell updates/sec

Title: US-09-778-516A-1

Sequence: 1 gatgtacgggccagatatatc.....gcgtgtcgcctagcgacgctc 8115

Scoring table:

	IDENTITY_NUC	Canopy 1 0
Canopy 10 0		

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-Exercise Metabolic Watch 09

Listing first 45 summaries

Database : N_Geneseq_032802:*

1	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT *
2	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT *
3	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT *
4	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT *
5	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT *
6	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT *
7	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT *
8	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT *
9	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT *
10	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT *
11	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT *
12	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT *
13	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT *
14	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT *
15	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT *
16	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT *
17	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT *
18	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT *
19	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT *
20	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT *
21	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT *
22	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT *
23	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT *
24	/SDSI/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT *

Pried. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3497.2	43.1	5059	21	AAZ36227	DNA encoding a pe
2	3110.6	38.3	3203	11	AAO06760	Sequence of beta-
3	1307.4	16.1	11592	22	AAEF83089	Elav expression v
4	1455.2	15.3	3853	19	AAV40006	Plasmid pCRM. Ch
5	1152.4	14.2	5306	21	AAVS3878	Expression vector
6	1148.2	14.1	14455	19	AAV32374	Complete sequence
7	1148.2	14.1	14455	21	AAAS9050	Nucleotide sequen
8	1148.2	14.1	14455	24	ABR94259	Nucleotide sequen
9	1140	14.0	5377	21	AAAS3872	Expression vector

[illegible]

ALIGNMENTS

RESULT	1
AA236227	ID
AA36227	standard; DNA; 5059 BP.
XX	AC
XX	AA36227;
XX	22-FEB-2000 (first entry)
XX	DE
XX	DNA encoding a permease an beta-galactosidase
XX	XX
KW	Lactose operon; permease; beta-galactosidase;
KW	beta-galactosidase activity; lactose fermenta
KW	fermented milk product; Streptococcus thermophilus
XX	XX
OS	Lactobacillus bulgaricus.
XX	XX
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	122..1873
FT	/tag= a
FT	/product= "permease"
FT	/note= "encodes AAY53620"
FT	1877..4519
FT	/tag= b
FT	/product= "beta-galactosidase"
FT	/note= "encodes AAY53621"
XX	XX
PN	FR27/8921-A1.
XX	XX
PD	26-NOV-1999.
XX	XX
FE	22-MAY-1998;
XX	98FR-0006456.

Expression vector	plasmid pCDNA3. C
Renilla luciferase	plasmid pRL-TK
Expression vector	plasmid pCTM1. Ch
Vector pshuttle DN	Novel plasmid DNA
Novel plasmid DNA	Novel plasmid DNA
Vaccine vector pM2	PMR9/SEO-5 vaccine
PMR9/SEO-5 vaccine	PMR9/1077 vector DN
Vector plasmid pCM	HCV NS3A ORF comp
HCV NS3A ORF comp	Nucleotide sequence
plasmid pCMV11. C	plasmid pCMV11. C
plasmid pCMV11. C	plasmid pCMV11-PS2
plasmid pCMV11 opt	plasmid pCMV11 opt
plasmid pCMV11-E2	HCV NS3 ORF compr
HCV NS3 ORF compr	plasmid pCR(TM)1a
plasmid pCR(TM)1a	AAV vector sequence
AAV vector sequence	AAV vector sequence
AAV vector sequence	Nucleotide sequence
Nucleotide sequence	Nucleotide sequence
Nucleotide sequence	Nucleotide sequence
Nucleotide sequence	DNA encoding hGH, h
DNA encoding hGH, h	Epidermal growth f
Epidermal growth f	plasmid pMRG1630 n
plasmid pMRG1630 n	Human alpha-L-idur
Human alpha-L-idur	Extracellular express

PR 22-MAY-1998; 98FR-0006456.
XX
PA (DANO-) CIE DANONE SA GERVAIS.
XX
PI Benbadis L, Brignon P, Genre F;
XX WPI; 2000-026146/03.
DR P-PSDB; AAY53620, AAY53621.
XX
PT New mutant *Lactobacillus bulgaricus* strain lacking beta-galactosidase
PT activity useful for making fermented milk products, especially yoghurt
PT
XX
PS Disclosure; Page 15-22; 32pp; French.
XX
CC The present sequence represents a lactose operon of *Lactobacillus*
CC *bulgaricus*, and encodes a permease protein and beta-galactosidase
CC protein. The specification describes mutant *L. bulgaricus* strain lacking
CC beta-galactosidase activity, because it has a termination (nonsense)
CC mutation in at least one of the coding sequences of the lactose operon.
CC The resulting mutant strain is unable to ferment lactose. Since the
CC mutant strain cannot ferment lactose, fermented milk products produced
CC using it do not have to be cooled to prevent post-acidification. The
CC mutant *L. bulgaricus* strain of the invention is used for producing a
CC fermented milk product, especially yoghurt. Especially, milk is
CC fermented with a culture of the mutant strain and optionally
CC *Streptococcus thermophilus* in the presence of at least one sugar
CC assimilable by the mutant strain, preferably glucose.
XX
SO Sequence 5059 BP; 1229 A; 1341 C; 1323 G; 1166 T; 0 other;

Query Match 43.1%; Score 3497.2; DB 21; Length 5059;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 3518; Conservative 0; Mismatches 13; Indels 3; Gaps 1;
QY 4584 ccgataaaattgagcgacggttggccctggtgacgacgacgagttaagttctacgac 4643
DB 1479 ccgaataaaattgagcgacggttggccctggtgacgacgacgagttaagttctacgac 1538
QY 4644 cattgcgcgtactctgcgcacgctgacgaagcccgacactcgtctcctggaattg 4703
DB 1539 cattgcgcgtactctgcgcacgctgacgaagcccgacactcgtctcctggaattg 1598
QY 4704 aacatggtgcttgcgtctgattcaacctggtgacgacggtcaaatlaaacygagctg 4763
DB 1599 aacatggtgcttgcgtctgattcaacctggtgacgacggtcaaatlaaacygagctg 1658
QY 4764 gcttgcgtactatgtttaaagagcgacgacggttagaagccgacgacgacgtcctggaat 4823
DB 1659 gcttgcgtactatgtttaaagagcgacgacgacggttagaagccgacgacgacgtcctggaat 1718
QY 4824 tctggagcccgcgatcgaagcgaagctggaagcgaacggttaactgacggtcatca 4883
DB 1719 tctggagcccgcgatcgaagcgaagctggaagcgaacggttaactgacggtcatca 1778
QY 4884 acagcgaaacttccaaatagccagatgctcttgcgacgtccacagcgctccaaagcc 4943
DB 1779 acagcgaaacttccaaatagccagatgctcttgcgacgtccacagcgctccaaagcc 1838
QY 4944 tggatgatgatattcaagttagaaggaagaatttagaataagcacaataagttagtaaaag 5003
DB 1839 tggatgatgatattcaagttagaaggaagaatttagaataagcacaataagttagtaaaag 1898
QY 5004 aaaaaaggttgacacgacgacgtgctgctgactgacgacccggaatttaccgaatca 5063
DB 1899 aaaaaaggttgacacgacgacgtgctgctgactgacgacccggaatttaccgaatca 1958
QY 5064 atacaattccccgcgacactcgacacatgaagctcttccaaagccaggaagaactggaagag 5123
DB 1959 atacaattccccgcgacactcgacacatgaagctcttccaaagccaggaagaactggaagag 2018
QY 5124 gcaagtcagtttagtcagtcctctgacgggagctgctgattactacgctgaacag 5183

DB 2019 gcaagtcagtttagtcagtcctctgacgggagctgctgattactacgctgaacag 2078
QY 5184 gccaggagacgactaaactctatcagaagaactttgacgacgaatttaagtcagtc 5243
DB 2079 gccaggagacgactaaactctatcagaagaactttgacgacgaatttaagtcagtc 2138
QY 5244 aagtaaccgcaacctggaacttgaaggtcttgcgcagcccaatgatgtcaacgtccaat 5303
DB 2139 aagtaaccgcaacctggaacttgaaggtcttgcgcagcccaatgatgtcaacgtccaat 2198
QY 5304 atccatggacgagtgagtagagatttccgcgcccaattccaaagcaaaatccgcgtcg 5363
DB 2199 atccatggacgagtgagtagagatttccgcgcccaattccaaagcaaaatccgcgtcg 2258
QY 5364 ctcttaagttaagtaactttgacgtgtagaagcttcttggaagaagaagtcagcttga 5423
DB 2259 ctcttaagttaagtaactttgacgtgtagaagcttcttggaagaagaagtcagcttga 2318
QY 5424 agttgacggggcgcaacacgcatctatgtctgtgtaacgacgacactcgtcgctacg 5483
DB 2319 agttgacggggcgcaacacgcatctatgtctgtgtaacgacgacactcgtcgctacg 2378
QY 5484 gggaagactccttaaccacagcgagtttagttaccgaattctctcaagaagaataa 5543
DB 2379 gggaagactccttaaccacagcgagtttagttaccgaattctctcaagaagaataa 2438
QY 5544 accgcctgcagctgctctccaaagatctctccgcctctgctgtaagaacagagact 5603
DB 2439 accgcctgcagctgctctccaaagatctctccgcctctgctgtaagaacagagact 2498
QY 5604 tctggcgacgtctggttcttcacagatcagtcacttccaaagccagcgtctgcaat 5663
DB 2499 tctggcgacgtctggttcttcacagatcagtcacttccaaagccagcgtctgcaat 2558
QY 5664 tggagacacttaagcttcgcgcacgcttgacgataacttaccaaaaggaagctggagag 5723
DB 2559 tggagacacttaagcttcgcgcacgcttgacgataacttaccaaaaggaagctggagag 2618
QY 5724 tcgaagcacaatctgctcactacgcttgccaatctgccaagctttaaagctggagagtgcgagata 5783
DB 2619 tcgaagcacaatctgctcactacgcttgccaatctgccaagctttaaagctggagagtgcgagata 2678
QY 5784 gtgaagtgacttggtgtgtctgtaaaagctggtgcccacatagaagcgaagcttggaattca 5843
DB 2679 gtgaagtgacttggtgtgtctgtaaaagctggtgcccacatagaagcgaagcttggaattca 2738
QY 5844 ctctggtcgtattgctcagtagctgctgtagcgcggaagaacctaactttaccaggtcc 5903
DB 2739 ctctggtcgtattgctcagtagctgctgtagcgcggaagaacctaactttaccaggtcc 2798
QY 5904 gctgtatttataccagcagcgacgctctttagaaggttagccggaagaagtgggtttcc 5963
DB 2799 gctgtatttataccagcagcgacgctctttagaaggttagccggaagaagtgggtttcc 2858
QY 5964 gcaactttgaactaaagaacggtatagatcttaacgaagcgcagcggtcttcttaag 6023
DB 2859 gcaactttgaactaaagaacggtatagatcttaacgaagcgcagcggtcttcttaag 2918
QY 6024 ggccaacccgcaacgaatttgaacagttggtgctggtgacataccggaagaagatga 6083
DB 2919 ggccaacccgcaacgaatttgaacagttggtgctggtgacataccggaagaagatga 2978
QY 6084 tctggacatacaagacatlgagcgaagcaacatcaatcaatgctgtccgctctcaatcc 6143
DB 2979 tctggacatacaagacatlgagcgaagcaacatcaatcaatgctgtccgctctcaatcc 3038
QY 6144 cgaacagtcctcttcttaaccggtctgtgacaagtacggccttaagtcattgaatga 6203
DB 3039 cgaacagtcctcttcttaaccggtctgtgacaagtacggccttaagtcattgaatga 3098
QY 6204 ctaacctggaagccacgacgaccttgggaaaagttggggggcacaagaatctctagcttca 6263

Db 3099 cttaacctggaagccacgcgacctgggaaaaagltggggggcgacgaagatcctagcttca 3158
OY 6264 atgtccagagcgatgacccaacatgtgtggaacccagcttatcccggtgtgaagaatga 6223
Db 3159 atgttccagagcgatgacccaacatgtgtggaacccagcttatcccggtgtgaagaatga 3218
OY 6324 tggctcggggaacaagaacatgcttcaatcctaaactgtgtctttaggaacaagtcttaag 6383
Db 3219 tggctcggggaacaagaacatgcttcaatcctaaactgtgtctttaggaacaagtcttaag 3278
OY 6384 ccggcaactgtcttgcgccaaatggtgtatagctccggaaagctgtatccgacccgggttc 6443
Db 3279 ccggcaactgtcttgcgccaaatggtgtatagctccggaaagctgtatccgacccgggttc 3338
OY 6444 agacatctgaaggggtgacccacaacgggaagtttagacggcgacccaattggaagcc 6503
Db 3339 agacatctgaaggggtgacccacaacgggaagtttagacggcgacccaattggaagcc 3398
OY 6504 ggaatgatgtcccgccgaaggtaatgtgaagaatacttgacccaataaaccggccaagccat 6563
Db 3399 ggaatgatgtcccgccgaaggtaatgtgaagaatacttgacccaataaaccggccaagccat 3458
OY 6564 ttatctcagttgataacgtctacgcgcacatgggcacatccgcgtgtgacctgtgcccgtaca 6623
Db 3459 ttatctcagttgataacgtctacgcgcacatgggcacatccgcgtgtgacctgtgcccgtaca 3518
OY 6624 cggcccttgaaaaataatccccactaccagggcggtctcaactgtgagctgtgattgaccaa 6683
Db 3519 cggcccttgaaaaataatccccactaccagggcggtctcaactgtgagctgtgattgaccaa 3578
OY 6684 gactggaaaaaagacggcgacactgtcttatgtggggcgactcgatgacggcgccaacggact 6743
Db 3579 gactggaaaaaagacggcgacactgtcttatgtggggcgactcgatgacggcgccaacggact 3638
OY 6744 atgaattctgggggaacggcgctgtgtctgtgtgacggactgaatccgcgaacgtgcta 6803
Db 3639 atgaattctgggggaacggcgctgtgtctgtgtgacggactgaatccgcgaacgtgcta 3698
OY 6804 atgtcaagcgctttagcccaacttaagttagaagtaaaagatgggcaacttctctca 6863
Db 3699 atgtcaagcgctttagcccaacttaagttagaagtaaaagatgggcaacttctctca 3758
OY 6864 aaaaacgaataattattacccaacagctcatcttactactctgtgactagcttctgtcg 6923
Db 3759 aaaaacgaataattattacccaacagctcatcttactactctgtgactagcttctgtcg 3818
OY 6924 atgtgcaagtgtgacctaccagaagcgcgctctgacacttggctcggaagctgtgacgaatccg 6983
Db 3819 atgtgcaagtgtgacctaccagaagcgcgctctgacacttggctcggaagctgtgacgaatccg 3878
OY 6984 ggaacttgcgcctgtgtgacgggaagtgcgtatgaaaaaggggaggtctgtctaccggg 7043
Db 3879 ggaacttgcgcctgtgtgacgggaagtgcgtatgaaaaaggggaggtctgtctaccggg 3938
OY 7044 taacggcccaactaaagaagacttgctctgtgacggatgagggcttcaactgtgtcgtgaag 7103
Db 3939 taacggcccaactaaagaagacttgctctgtgacggatgagggcttcaactgtgtcgtgaag 3998
OY 7104 cagaagaagtagtctcaaaagctgtgcgggaatttagccgggaagggcgccgaatttagtctg 7163
Db 3999 cagaagaagtagtctcaaaagctgtgcgggaatttagccgggaagggcgccgaatttagtctg 4058
OY 7164 atccgaactacaacctagagcctgaagaaggaataactccaattctcttcccaagtca 7223
Db 4059 atccgaactacaacctagagcctgaagaaggaataactccaattctcttcccaagtca 4118
OY 7224 aaggcttgcccggttccctcaaglatgcccgttaggggaatacttgaagcggctgcgggaat 7283
Db 4119 aaggcttgcccggttccctcaaglatgcccgttaggggaatacttgaagcggctgcgggaat 4178
OY 7284 ttacactctggcgcccttgacgacgaacgaacggggaagtggtttacgctatgactctg 7343
Db 4179 ttacactctggcgcccttgacgacgaacgaacggggaagtggtttacgctatgactctg 4238

OY 7344 cccggtgggaaaaatgcccggcaagtatgcccgtctggaagaacatcaagctgcggaagtaag 7403
Db 4239 cccggtgggaaaaatgcccggcaagtatgcccgtctggaagaacatcaagctgcggaagtaag 4298
OY 7404 aagaactcgttctgtgtcaagaactgcttcaactgtgtcctgtgccttcaaaaggtgatttaa 7463
Db 4299 aagaactcgttctgtgtcaagaactgcttcaactgtgtcctgtgccttcaaaaggtgatttaa 4358
OY 7464 ccgtgacctatgaagtcgaatggaacgggggaagattgctgtcaaaagctgacttcccaggcg 7523
Db 4359 ctgtgacctatgaagtcgaatggaacgggggaagattgctgtcaaaagctgacttcccaggcg 4418
OY 7524 cgggaagaagctgtgtctcttccagccttggctgtgaacccggcgcccgccgaagaactga 7583
Db 4419 cgggaagaagcgtgtctcttccagccttggctgtgaacccggcgcccgccgaagaactga 4478
OY 7584 ccgattacccgtactatgctgtctggacctaaatgagaagctaacccaacggcttggaaagta 7643
Db 4479 ccgattacccgtactatgctgtctggacctaaatgagaagctaacccaacggcttggaaagta 4538
OY 7644 attacttgggcatctaccagggcggttaaaaaagaacttaagccata---tcgtccgc 7700
Db 4539 attacttgggcatctaccagggcggttaaaaaagaacttaagccatactgtgtccgc 4598
OY 7701 aggaaaacgggcaacccggaagaagtttcgctgtgtacacgctcttgaatgaagaaggcggt 7760
Db 4599 aggaaaacgggcaacccggaagaagtttcgctgtgtacacgctcttgaatgaagaaggcggt 4658
OY 7761 tggaaattacgccaatgvgggcagacttgaactgtgtctgttgcatacttctgcgcgc 7820
Db 4659 tggaaattacgccaatgvgggcagacttgaactgtgtctgttgcatacttctgcgcgc 4718
OY 7821 aaattgaagcagcggaacagcttcttgaactgaactaaacattacatttgggttaaggcct 7880
Db 4719 aaattgaagcagcggaacagcttcttgaactgaactaaacattacatttgggttaaggcct 4778
OY 7881 taagcccccagatctgggggtctgcgcgggatgactccttgggggcaggaaggtccacccggaat 7940
Db 4779 taagcccccagatctgggggtctgcgcgggatgactccttgggggcaggaaggtccacccggaat 4838
OY 7941 tctgcctgtgagtctcaaaaagcccgccagcttcgcctgtgtgatttcagcccccttactaa 8000
Db 4839 tctgcctgtgagtctcaaaaagcccgccagcttcgcctgtgtgatttcagcccccttactaa 4898
OY 8001 aataaatgtctacattgacttaacagatgaatatttagtaaaagaagaacggttagga 8060
Db 4899 aataaatgtctacattgacttaacagatgaatatttagtaaaagaagaacggttagga 4958
OY 8061 agatgccaacgatacagaagaagtgcacaagcagccggtgtctgctagcgacggt 8114
Db 4959 agatgccaacgatacagaagaagtgcacaagcagccggtgtctgctagcgacggt 5012

RESULT 2
AA006760
ID AA006760 standard; DNA; 3203 BP.
XX
AC AA006760;
XX
DT 25-FEB-1991 (first entry)
XX
DE Sequence of beta-galactosidase gene.
XX
KM pBR329; pBG1; ds.
XX
OS Lactobacillus bulgaricus SBT0034.
XX
FH Key Location/Qualifiers
FT 285..3083
FT CDS /*tag= a
XX
PN JP02261383-A.

XX 24-OCT-1990.
PD 31-MAR-1989; 89JP-0082851.
XX 31-MAR-1989; 89JP-0082851.
PR 31-MAR-1989; 89JP-0082851.
XX (SNOW) SNOW BRAND MILK PRODUCTS.
XX WPI: 1990-364398/49.
DR P-PSDB; AAR08132.
XX
PT DNA from *Lactobacillus bulgaricus* SIB0034 encoding beta-
PT galactosidase - used to produce high yields of beta-galactosidase by
PT recombinant DNA technology.
XX
PS Claim 1; Fig 1; 6pp; Japanese.
CC Beta-galactosidase may be produced from an *E.coli* expression system
CC transformed by the gene sequence incorporated in the plasmid vector
CC pBel, derived from pBR329.
XX
SQ Sequence 3203 BP; 836 A; 807 C; 867 G; 693 T; 0 other;

Query Match 38.3%; Score 3110.6; DB 11; Length 3203;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 3177; Conservative 0; Mismatches 14; Indels 8; Gaps 5;

QY 4920 cgatcgccagcagcgtcccaagccctgtgagtgtatctatcaagttgaaggaagaattaga 4979
D 1 cgatcgccagcagcgtcccaagccctgtgagtgtatctatcaagttgaaggaagaattaga 60
QY 4980 aaatgagcaataagttagtaaaagaaaaaagagttgaccagcgaccctgtgctgtga 5039
D 61 aaatgagcaataagttagtaaaagaaaaaagagttgaccagcgaccctgtgctgtga 120
QY 5040 ctgaccgggaagtttaagcaatcaataaattcccccgcacatcgacatgaatcttcc 5099
D 121 ctgaccgggaagtttaagcaatcaataaattcccccgcacatcgacatgaatcttcc 180
QY 5100 aagcccggaagaaatgaggaaggaaggaagtcacatttgaatgacatccctggaaggaat 5159
D 181 aagcccggaagaaatgaggaaggaaggaagtcacatttgaatgacatccctggaaggaat 239
QY 5160 ggcctgattgacgtctgaaacacggccagggaccagtcacatctctatgcaagaacttg 5219
D 240 ggcctgattgacgtctgaaacacggccagggaccagtcacatctctatgcaagaacttg 299
QY 5220 acgataagcaatttaagtcaatcaaatgtaacccggcaacctgtgaactgcaaggtttgac 5279
D 300 acgataagcaatttaagtcaatcaaatgtaacccggcaacctgtgaactgcaaggtttgac 359
QY 5280 agccccaatgtaacgttcaatcaatcaatcaatggaaggaaggaaggaaggaaggaagga 5339
D 360 agccccaatgtaacgttcaatcaatcaatcaatggaaggaaggaaggaaggaaggaagga 419
QY 5340 aaattccaagcaaaatccgtctctctatgtaagtaacttgaactgtgaactgtgaactgt 5399
D 420 aaattccaagcaaaatccgtctctctatgtaagtaacttgaactgtgaactgtgaactgt 479
QY 5400 tct-gggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 5458
D 480 tctggggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 539
QY 5459 ctgagagggcaactctgtggtgtaaggaaggaaggaaggaaggaaggaaggaaggaagga 5518
D 540 ctgagagggcaactctgtggtgtaaggaaggaaggaaggaaggaaggaaggaaggaagga 599
QY 5519 accaagttccccaagaaagaaataaacacgctgtgcaagtgtgctctcaagaatattctcc 5578
D 600 accaagttccccaagaaagaaataaacacgctgtgcaagtgtgctctcaagaatattctcc 659

QY 5579 gctcctgtgctggaagacaggaactctgtgcgcatgtctgtgttctgaagatcaagtact 5638
D 660 gctcctgtgctggaagacaggaactctgtgcgcatgtctgtgttctgaagatcaagtact 719
QY 5639 cttaagcccaagccgctgtcacttgaaggaacttaagcttaagcccaagcttaagcttaagct 5698
D 720 cttaagcccaagccgctgtcacttgaaggaacttaagcttaagcccaagcttaagcttaagct 779
QY 5699 aactaccaaaaaggaagcttgaagctgaagcccaatattgcttaacgcttgcgaatgccc 5758
D 780 aactaccaaaaaggaagcttgaagctgaagcccaatattgcttaacgcttgcgaatgccc 839
QY 5759 agctttaagctggaaggtgctggaaggaaggaaggaaggaaggaaggaaggaaggaagga 5818
D 840 agctttaagctggaaggtgctggaaggaaggaaggaaggaaggaaggaaggaaggaagga 899
QY 5819 atcagaagcgaagcagcaggaatctacacttgcctgtgatttgaagcagctgtgagcgcg 5878
D 900 atcagaagcgaagcagcaggaatctacacttgcctgtgatttgaagcagctgtgagcgcg 959
QY 5879 gaaaagcctaacttaccaggttcgcctgtattataccagcagcagcagcagcagcagcagc 5938
D 960 gaaaagcctaacttaccaggttcgcctgtattataccagcagcagcagcagcagcagcagc 1019
QY 5939 gttagccgcaaggaaggtgtgttccgcgaacttgaactaaagcaggaattatgactt 5998
D 1020 gttagccgcaaggaaggtgtgttccgcgaacttgaactaaagcaggaattatgactt 1079
QY 5999 aacgcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 6058
D 1080 aacgcaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1139
QY 6059 cgggctacacggaaggaagatataatctgtgacatcaagacatgaagcgaagcaacatc 6118
D 1140 cgggctacacggaaggaagatataatctgtgacatcaagacatgaagcgaagcaacatc 1199
QY 6119 aatgtctgcgctgtctcctacatccgaacacagctccttcttcggtctgtgacaag 6178
D 1200 aatgtctgcgctgtctcctacatccgaacacagctccttcttcggtctgtgacaag 1259
QY 6179 taaggccttaacgttaattgaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 6238
D 1260 taaggccttaacgttaattgaaggaaggaaggaaggaaggaaggaaggaaggaaggaagga 1319
QY 6239 ggggggacgaagaatcctagcttcaatgttccaaagcgaatgaacgaatgtgtggaagcc 6298
D 1320 ggggggacgaagaatcctagcttcaatgttccaaagcgaatgtgtggaagcc 1379
QY 6299 agcttatccgggtggaagaataatgagtgtcgggacaaagcaatcgttcaatcctaatc 6358
D 1380 agcttatccgggtggaagaataatgagtgtcgggacaaagcaatcgttcaatcctaatc 1439
QY 6359 tggctcttaagcaatgaatcttaacgcaagcagcttcttgcgaatgaagctgaatcagc 6418
D 1440 tggctcttaagcaatgaatcttaacgcaagcagcttcttgcgaatgaagctgaatcagc 1459
QY 6419 cgggaagcgtgctccgaacccgggttcaagcactaagaaggtgacccaacccggaaggtt 6478
D 1500 cgggaagcgtgctccgaacccgggttcaagcactaagaaggtgacccaacccggaaggtt 1559
QY 6479 gacgaagcacaacagatgaaggaagcaggaatgatactccggcaaggaatgaagaatc 6538
D 1560 gacgaagcacaacagatgaaggaagcaggaatgatactccggcaaggaatgaagaatc 1619
QY 6539 ttgaccaataaacagccaacgcaattatctcagttgaatcagctcagcagcagcagcagc 6598
D 1620 ttgaccaataaacagccaacgcaattatctcagttgaatcagctcagcagcagcagcagc 1678
QY 6599 tccgtctgtgacgtgtgccccttacaagcccttgaagaataatccccacatcagcagcagc 6658
D 1679 tccgtctgtgacgtgtgccccttacaagcccttgaagaataatccccacatcagcagcagc 1736
QY 6659 ttcatctgtgacgtgtgacacgaaggaaggaaggaaggaaggaaggaaggaaggaagga 6718

Dh	1737	ttcatccgggagcttggatcttgcacaaagcacttggaaaagacggtgcaccttcttaattgggggc	1736
Qy	6719	gacttcgatatccggcccaacccgaacttgaattcttggggaaaacggtctgtcttcttgcatac	6778
Dh	1797	gacttcgatatccggcccaacccgaacttgaattcttggggaaaacggtctgtcttcttgcatac	1856
Qy	6779	cgagcttgaatccggccgaacacttgcataatgtccaaagccctttaaagcccttaagattgaag	6838
Dh	1857	cgagcttgaatccggccgaacacttgcataatgtccaaagccctttaaagccctttaaagattgaag	1916
Qy	6839	gtaaaagatbtggcagcccttccctcccaaaaacgcaattttatccaaagcgtcacattcac	6898
Dh	1917	gtaaaagatbtggcagcccttccctcccaaaaacgcaattttatccaaagcgtcacattcac	1976
Qy	6899	tactcttgaactagctcttcttgttgcataatggcgaattgacattcaagacgagcccttgcac	6958
Dh	1977	tactcttgaactagctcttcttgttgcataatggcgaattgacattcaagacgagcccttgcac	2036
Qy	6959	tttggcccttggagcccttggccggaatccggagacctttgcaccttgccttcttggccgggaagttgcatt	7018
Dh	2037	tttggcccttggagcccttggccggaatccggagacctttgcaccttgccttcttggccgggaagttgcatt	2096
Qy	7019	gaaaagaggggaggttcgtctctacccgggttgaacggcccaacttaaaagagagacttcccttggcg	7078
Dh	2097	gaaaagaggggaggttcgtctctacccgggttgaacggcccaacttaaaagagagacttcccttggcg	2156
Qy	7079	gatgaggggcttcaacttgtgtcttgaagacgaagaagtacgtctcaaaagctgcgcgggaatttaag	7138
Dh	2157	gatgaggggcttcaacttgtgtcttgaagacgaagaagtacgtctcaaaagctgcgcgggaatttaag	2216
Qy	7139	ccggaaagggcgccgaagatattagttctgaatcccgactaaacacttgggctttaaaggaataaac	7198
Dh	2217	ccggaaagggcgccgaagatattagttctgaatcccgactaaacacttgggctttaaaggaataaac	2276
Qy	7199	ttccaaattctctcttcccaaagtccaagggcttgcgggttccctcaagtatatgcgcgttag	7258
Dh	2277	ttccaaattctctcttcccaaaggtccaagggcttgcgggttccctcaagtatatgcgcgttag	2336
Qy	7259	gaatacatttgaagcggctgcgcgggaatttacctctcttgcgggccccttgaacggaacacggcg	7318
Dh	2337	gaatacatttgaagcggctgcgcgggaatttacctctcttgcgggccccttgaacggaacacggcg	2396
Qy	7319	ggagcttggctaaagctatgaatctgcgcccgggttgggaaaatctgcggcggaatctatgcgcgcttgc	7378
Dh	2397	ggagcttggctaaagctatgaatctgcgcccgggttgggaaaatctgcggcggaatctatgcgcgcttgc	2456
Qy	7379	aaaagaactcagcgtcggcgaggtccaaggaagagacctccggttttgcataaagcgtccttaagctt	7438
Dh	2457	aaaagaactcagcgtcggcgaggtccaaggaagagacctccggttttgcataaagcgtccttaagctt	2516
Qy	7439	cccttgcgcttcaaaggggttgaatttcaaacggttgacacttgaagttcgaatgcggcggaagatt	7498
Dh	2517	cccttgcgcttcaaaggggttgaatttcaaacggttgacacttgaagttcgaatgcggcggaagatt	2576
Qy	7499	gctgttaacagcttgaacttcccaaggcgcggaagaagcttgccttcttgcacagccttggcttgc	7558
Dh	2577	gctgttaacagcttgaacttcccaaggcgcggaagaagcttgccttcttgcacagccttggcttgc	2636
Qy	7559	aaacttggcccttgcgaaaaagacgtacacggattaccgttactatgtgtccttgggaccttaatgag	7618
Dh	2637	aaacttggcccttgcgaaaaagacgtacacggattaccgttactatgtgtccttgggaccttaatgag	2696
Qy	7619	agctaccagaacgcttcttgaaggttaattacacttgggcatctacacagggsgcgttaaaaaag	7678
Dh	2697	agctaccagaacgcttcttgaaggttaattacacttgggcatctacacagggsgcgttaaaaaag	2756
Qy	7679	aaacttggcccatatccttgcgtcccgcaaggaacggcggaacgggaagatctgcgttgcatac	7735
Dh	2757	aaacttggcccatatccttgcgtcccgcaaggaacggcggaacgggaagatctgcgttgcatac	2816
Qy	7736	caagcttcttgaataaagggcgcttcttgaattcttgcgaacatttggggcagacttgaacttgc	7795

Db	2817	cagctctttgatgaaaggcgcgcttggatattcagccaaatggggcagacttgaactg	2876		
Qy	7796	tctgctttgcacatactctgcgcccaaatgtgaagcagcgagccacgctttgaactgact	7855		
Db	2877	tctgctttgcacatactctgcgcccaaatgtgaagcagcgagccacgctttgaactgact	2936		
Qy	7856	aacattacactttggtttagagccttaagcgcccaagtgggggtcggcggttagctcc	7915		
Db	2937	aacattacactttggtttagagccttaagcgcccaagtgggggtcggcggttagctcc	2996		
Qy	7916	tgggggcagaaggtgcccccgaattctgtcctgtgattctcctaaagcccgccagctcgc	7975		
Db	2997	tgggggcagaaggtgcccccgaattctgtcctgtgattctcctaaagcccgccagctcgc	3056		
Qy	7976	ctggtgatctcagcccttctaactaaatgtacaaatgtacttcaacagatgaatt	8035		
Db	3057	ctggtgatctcagcccttctaactaaatgtacaaatgtacttcaacagatgaatt	3116		
Qy	8036	ttagtaaaagcaagcgagtgagaagatgtgcaacgatacagagaagtgcgaagcgacgcg	8095		
Db	3117	ttagtaaaagcaagcgagtgagaagatgtgcaacgatacagagaagtgcgaagcgacgcg	3176		
Qy	8096	gcgtgtcgtcgtacgacgcgt	8114		
Db	3177	gcgtgtcgtcgtacgacgcgt	3195		
RESULT 3					
ID	AAFB83089				
ID	AAFB83089	standard; DNA; 11592 BP.			
AC	AAFB83089;				
XX	29-JUN-2001	(first entry)			
DT					
XX	E1AV expression vector pONT8za CMVhyb sequence.				
DE					
XX	Retrovirus; recombinase recognition sequence; RRS; LTR; recombinase;				
KW	long terminal repeat; pharmaceutical; cytosolic; antiinflammatory;				
KW	antirheumatic; antiarthritic; antialsthmatic; osteopathic; cardiant; E1AV				
KW	vasotropic; neuroprotective; nootropic; cerebroprotective; antipsoriatic				
KW	antiarteriosclerotic; vulnerrary; anti-HIV; antilucer; thrombolytic;				
KM	dermatological; gene therapy; ss.				
XX					
OS	Synthetic.				
OS	Equine infectious anemia virus.				
XX					
PN	WO200125466-A1.				
XX					
PD	12-APR-2001.				
XX					
PF	05-OCT-2000; 2000WO-GH03837.				
XX					
PR	05-OCT-1999; 99GB-0023558.				
XX					
PA	(OXFO-) OXFORD BIOMEDICA UK LTD.				
XX					
PI	Slingsby J, Kingsman SM, Rohll J, Slade A;				
XX					
DR	WPI; 2001-281732/29.				
XX					
PT	Modifying producer cells for making retrovirus by transfecting with a				
PT	construct comprising a 5'-recombinase recognition sequence, long				
PT	terminal repeat and 3'-recombinase recognition sequence, in presence				
PT	of recombinase -				
XX					
XX	Examples; Page 111-114; 133pp; English.				
XX					
CC	The invention relates to a method of modifying producer cells for making				
CC	retrovirus by transfecting with a construct comprising a 5'-recombinase				
CC	recognition sequence (RRS), long terminal repeat (LTR) and 3'-RRS, in				
CC	presence of recombinase. The regulated retroviral vector produced is				
CC	useful in the manufacture of a pharmaceutical composition to deliver a				

[illegible]

Qy	1288	accgtaaaagccgcgltgcgtgcggtttttccatagctccgccccctgaagcagc	1347
Db	2070	accgtaaaagccgcgltgcgtgcggtttttccatagctccgccccctgaagcagc	2129
Qy	1348	acaaaacgcagcgtctcaagtcagagtggtgagaaacccgcagagctataagatccag	1407
Db	2130	acaaaacgcagcgtctcaagtcagagtggtgagaaacccgcagagctataagatccag	2189
Qy	1408	cgttccccctggaagatccctcgtctgcgctctctctgttccgaacctgcgcttaacggat	1467
Db	2190	cgttccccctggaagatccctcgtctgcgctctctctgttccgaacctgcgcttaacggat	2249
Qy	1468	acctgcgcgcttctcccttcgggaaagcttgcgcttcttctaagtcacgctyagct	1527
Db	2250	acctgcgcgcttctcccttcgggaaagcttgcgcttcttctaagtcacgctyagct	2309
Qy	1528	atccagttcggtttaagtcgtttcgcgtccaaagcttgagctgtgtgcagaaacccccgttc	1587
Db	2310	atccagttcggtttaagtcgtttcgcgtccaaagcttgagctgtgtgcagaaacccccgttc	2365
Qy	1588	agcccgaccgctgcgccttalcgccgtaactatcgtcttgaftccaaacccggtaaagacag	1647
Db	2370	agcccgaccgctgcgccttalcgccgtaactatcgtcttgaftccaaacccggtaaagacag	2429
Qy	1648	actatgcgcactgycagcagccacgtgtaacaggtatgcagagcagagatgtatgagcg	1707
Db	2430	actatgcgcactgycagcagccacgtgtaacaggtatgcagagcagagatgtatgagcg	2489
Qy	1708	gtgctacagaafttctttaaattgtgtgacctaacacgcgtcacactagaagacagatttg	1767
Db	2490	gtgctacagagttctttaaagttgtgtgacctaacacgcgtcacactagaagacagatttg	2549
Qy	1768	gtatctgcgctctgcgtcgaagccagcttaaccttcggaaagaattggtagctcttgatccg	1827
Db	2550	gtatctgcgctctgcgtcgaagccagcttaaccttcggaaagaattggtagctcttgatccg	2609
Qy	1828	gcaaaacaaacccgcgctgtagcggtgtgttttttttgcaagaacagcagattacgcga	1887
Db	2610	gcaaaacaaacccgcgctgtagcggtgtgttttttttgcaagaacagcagattacgcga	2665
Qy	1888	gaaaaaaagatctctcaagaagaatcccttgatcttcttctacaggggtctgacgctcagtgga	1947
Db	2670	gaaaaaaagatctctcaagaagaatcccttgatcttcttctacaggggtctgacgctcagtgga	2729
Qy	1948	acgaaactcagctgaaggatttggtaatgagcggaatataattgaattattaga	2007
Db	2730	acgaaactcagctgaaggatttggtaatgagcggaatataaaagaatcttaccctaga	2789
RESULT	5		
AAAS3878			
ID	AAAS3878	standard; DNA; 5306 BP.	
AC	AAAS3878;		
DT	03-JAN-2001	(first entry)	
XX	Expression: vector pRiG-22.		
DE			
XX	Vector: endogenous gene; activation; over-expression;		
KM	erythropoietin; growth hormone; drug discovery;		
KW	granulocyte colony stimulating factor; ds.		
XX			
OS	Synthetic.		
XX			
PN	WO200049162-A2.		
PD	24-AUG-2000.		
XX			
PF	22-FEB-2000; 2000MO-US04429.		
XX			
PR	19-FEB-1999; 99US-0253022.		

PR 08-MAR-1999; 99US-0263814.
PR 26-MAR-1999; 99US-0276820.
XX
XX (ATHE-) ATHERSYS INC.
XX
XX Harrington JJ, Sherf B, Rundlett S;
XX
XX MPI; 2000-549276/50.
PT Non-targeted activation of endogenous genes, e.g. for the production of
PT erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT proteins and for drug discovery
XX
XX
PS Disclosure; Fig 35; 240pp; English.
CC
CC New methods, vectors and cells are described for non-targeted
CC activation and over-expression of endogenous genes. The vector
CC constructs comprise transcriptional regulatory sequences (TRS) and
CC unpaired splice donor sequences (USDS), preferably the vectors
CC comprise (in sequential order) a TRS, an USDS, a rare cutting
CC restriction site (RCRS) and a linearization site (LS) with a second
CC TRS linked to a selectable marker (SM) lacking a polyadenylation
CC signal. The methods, vectors and cells comprising the vectors may
CC be used for the non-targeted activation and over-expression of
CC endogenous genes, e.g. for the production of proteins (including
CC erythropoietin, growth hormone or granulocyte-colony stimulating
CC factor) and drug discovery. The advantage of these methods are that
CC endogenous genes including those associated with human disease and
CC development, may be activated and isolated without prior knowledge
CC of the sequence structure, function or expression profile of
CC the genes being known.
XX
XX
SQ Sequence 5306 BP; 1303 A; 1365 C; 1383 G; 1255 T; 0 other;

Query Match 14.2%; Score 1152.4; DB 21; Length 5306;
Best Local Similarity 76.4%; Pred. No. 4.3e-163;
Matches 1524; Conservative 0; Mismatches 396; Indels 74; Gaps 6;

QY 23 gtggaattgatatgacagtattatataagtaacatcaatgaaggtgacattgacata 82
DB 140 gtggaattgatatgacagtattatataagtaacatcaatgaaggtgacattgacata 199
QY 83 gccacataatgaggtccgcgtttacataacttaacgttaaatggccgcctgctgacgc 142
DB 200 gccacataatgaggtccgcgtttacataacttaacgttaaatggccgcctgctgacgc 259
QY 143 ccaacgaccccgccacattgacgtlcaataatgacgtatgttccacatagtaacgcaatag 202
DB 260 ccaacgaccccgccacattgacgtlcaataatgacgtatgttccacatagtaacgcaatag 319
QY 203 ggaattccattgacgtacatggtgtgacatttaacgttaactggccacttggcagtac 262
DB 320 ggaattccattgacgtacatggtgtgacatttaacgttaactggccacttggcagtac 379
QY 263 atcaagtgtacatagcgaagtaacgcccctattgacgtcaatgacgttaaatggccg 322
DB 380 atcaagtgtacatagcgaagtaacgcccctattgacgtcaatgacgttaaatggccg 439
QY 323 cctggaattatgccagtaacgtacttaatggaacttccacttggcagtaacatag 382
DB 440 cctggaattatgccagtaacgtacttaatggaacttccacttggcagtaacatag 499
QY 383 tatatgacatcgctattacatggtgtgacgttttggcagtaacatgagcggtgat 442
DB 500 tatatgacatcgctattacatggtgtgacgttttggcagtaacatgagcggtgat 559
QY 443 agcggttgaactacaggggattcccaagtcacaccccatgacgtcaatggagttgt 502
DB 560 agcggttgaactacaggggattcccaagtcacaccccatgacgtcaatggagttgt 619
QY 503 tttagcacaacaaatcaacgggacttccaaatgtcgttaacaact-----ccgcccc 554
|||||

DB 620 tttagcacaacaaatcaacgggacttccaaatgtcgttaacaactgcatccgcccc 679
QY 555 attgacgcaaatggcggtgagctgtgacgttgggaggtctatataagcagaactctgt 614
DB 680 gttagcgaacaaatggcggtgagctgtgacgttgggaggtctatataagcagaactctgt 739
QY 615 gtaactagaagaacccactgacttgcgttactatcgaaatataagactacatagga 674
DB 740 agtgaacgcgtcagatcatcacttaattctgacgaactactacgttaacgccaatagagcctcc 759
QY 675 gaaccaaagcttggtagccagctcgatccacatagtaacggccaggtgtgtgt-gaatt 733
DB 800 tgcagaactgtcttagtgacaactatcgattccacacattatagcggccaggttaatt 859
QY 734 ctgcagatataccatcaacacgtggccgcctcaggaatgcatctagaggccattctat 793
DB 860 gtaacacgactcatatgacgtacgtccgggagcagaacagcccccacatggctgagtaata 919
QY 794 agtgcacacctaagtctagagctcgtctgacgaactcgtgactgtgcttctagttgcagc 853
DB 920 cgactcactatagggcgacagcgttgagtaactcgtactacttaagagag----- 965
QY 854 catctgtgttggccctccctcccgctgcttcccttgacccttggaaggtgcatcccaactg 913
DB 966 -----gctatctgcccagtagcagtcgaagaagaagat 1000
QY 914 tcccttccataataatgagaatgtgacgtatgtctgagtagtgcattctatc 973
DB 1001 ttaagaaagccgaacaaacagcgtcaatgaagccgaagtggcgaagccgattcccatcg 1060
QY 974 tgggggttgggttggggtgcaagacagcaaggggaggaatgttggaagaacatagcagcatg 1033
DB 1061 gtatgttcgagcatatagagccagcaacccactgtgtgcgcgtgtgagtcgcccagcag 1120
QY 1034 ctgggagatcggttgggtctctatgtgtcttgaagcggaacacacagctgcatatgaat 1093
DB 1121 atgcg-----ccgcgttagagagatccacagagcagtggtgtgtgcacatga 1166
QY 1094 cggccaaacgcggggagagagcggttgcgtatctggcgctctccgcttccctgcac 1153
DB 1167 tcgcgtatgtcatagtagtgcgtccaaagtagcgaagcgaagcagacttggcgcgccaaagc 1226
QY 1154 tgaactcgtcgtcgtcgttgcgttcggtcggcgagcgtatccgtcaactcaacgaagcgt 1213
DB 1227 ggtcgcgaagtagtctcg-agaaacggtgtgcacatagaatgtgcatcaacgatatagcgt 1285
QY 1214 aatacgttataccacgaatcaaggagataacgcaggaagaacatgtgacaagaagcca 1273
DB 1286 agatcccttgtagtagtcgagatctgtcgaagc-----catgtgacaaagcca 1334
QY 1274 gcaaaagccaggaacccgttaaaagccggttgcgtttttccataggtctcgccc 1333
DB 1335 gcaaaagccaggaacccgttaaaagccggttgcgtttttccataggtctcgccc 1394
QY 1334 cccctgaacgacatcacaaataatcgacgctcaagtcagaggttggcgaacacccgacagact 1393
DB 1395 cccctgaacgacatcacaaataatcgacgctcaagtcagaggttggcgaacacccgacagact 1454
QY 1394 ataaagatacagaggttcccccctggaagctccctcgtgtgcgtctcctgtttccgaacct 1453
DB 1455 ataaagatacagaggttcccccctggaagctccctcgtgtgcgtctcctgtttccgaacct 1514
QY 1454 ggcgcttacccgatacctgtcgcgttcccttccttcgggaaggttggcgcttccataag 1513
DB 1515 ggcgcttacccgatacctgtcgcgttcccttccttcgggaaggttggcgcttccataag 1574
QY 1514 ctcaacgtgtagatcatcgaatcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 1573
DB 1575 ctcaacgtgtagatcatcgaatcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgttgcgt 1634
QY 1574 cgaaccccccgttcaagcccgacgcgtgcgttccatccgttaactatcgtttagtccaa 1633
DB 1635 cgaaccccccgttcaagcccgacgcgtgcgttccatccgttaactatcgtttagtccaa 1694

```
QY 1634 ccggttaagacacactatcgccactgagcagccacgtgtacagattagcaagc 1693
    |||||||
Db 1695 ccggttaagacacactatcgccactgagcagccacgtgtacagattagcaagc 1754
QY 1694 gaggatgtagggcggtctacagagcttcgaagtgtgacctaacaggtacactag 1733
    |||||||
Db 1755 gaggatgtagggcggtctacagagcttcgaagtgtgacctaacaggtacactag 1814
QY 1754 aagacagattgtgtatctgcgtctgctgaagccagttacacttggaataaagttgg 1813
    |||||||
Db 1815 aagacagattgtgtatctgcgtctgctgaagccagttacacttggaataaagttgg 1874
QY 1814 tagctctgacccggaacaaacacacgcgtgtagcggtgtttttgtgttcaagca 1873
    |||||||
Db 1875 tagctctgacccggaacaaacacacgcgtgtagcggtgtttttgtgttcaagca 1934
QY 1874 gcaatatacggcagcaaaaaaagatctcaagaagatccttcttcttcaagggtc 1933
    |||||||
Db 1935 gcaatatacggcagcaaaaaaagatctcaagaagatccttcttcttcaagggtc 1994
QY 1934 tgacgtcagtgagcaaaaacacacgttaagattgtgtcaatgaagcgatacatatt 1993
    |||||||
Db 1995 tgacgtcagtgagcaaaaacacacgttaagattgtgtcaatgaagcgatacatatt 2054
QY 1994 tgaatgtattaga 2007
    |||||
Db 2055 gatctcacactaga 2068

RESULT 6
AAV32374/c
ID AAV32374 standard; DNA; 14455 BP.
XX
AC AAV32374;
XX
DT 29-SEP-1998 (first entry)
DE Complete sequence of the pEI/Fiber plasmid.
XX
KW Circular; adenovirus type 5; pEI/Fiber plasmid; structural protein;
KM complementation; fiber protein; gene therapy; HIV; tumour; AD5;
KW early gene; Huntington's disease; Ray-Sachs disease;
XX sickle cell disease; El regulatory protein; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 1460..4998
FT /tag= a
FT /note= "AD5 El regulatory gene"
FT misc_feature complement (10922..14233)
FT /tag= b
FT /note= "AD5 fiber gene consisting of a CMV
FT promoter at 5' end of this gene"
XX
XX W09813499-A2.
XX
XX 02-APR-1998.
XX
XX 24-SEP-1997; 97WO-EP05251.
XX
XX 25-SEP-1996; 96US-0719806.
XX
XX (NOVS ) NOVARTIS AG.
XX (SCRI ) SCRIPPS RES INST.
XX
XX Memerow GR, Von Seggern DJ;
XX
XX WPI; 1998-230709/20.
XX
XX Adenoviral vectors - which lack DNA encoding for structural protein
PT or fibre protein used particularly for gene therapy
```

```
XX
PS Example 1; Pages 112-131; 170pp; English.
XX
CC The present sequence is that of a pEI/Fiber plasmid used in the
CC method of the invention. The plasmid contains an adenovirus type 5 (AD5)
CC fiber gene controlled by a CMV promoter, an AD5 El gene and a pMAV
CC backbone. The invention provides adenoviral vectors having deletions of
CC all or part of various gene sequences encoding adenoviral structural
CC proteins and/or early region proteins. Deletions in these proteins would
CC allow a reduced risk of wild-type virus contamination and would also
CC allow packaging of foreign DNA in such vectors for a variety of
CC diagnostic and therapeutic applications. The adenoviral vectors having
CC deletions in the structural and/or early gene regions are produced by
CC cellular complementation of these adenoviral genes. Therefore, the
CC pEI/Fiber plasmid was used as a complementation plasmid which was
CC introduced into a host cell line where parts of the fiber and El gene
CC region would be stably inserted into the host cell chromosomes. The
CC resulting El/fiber gene deficient plasmid can be used as a gene delivery
CC vector. The vectors can be used for diagnosis or gene therapy, e.g. for
CC treating conditions characterised by hyper-proliferative cells (e.g.
CC tumours), genetic diseases (e.g. Huntington's disease, Tay-Sachs disease,
CC or sickle cell disease), or infections (e.g. HIV infection). They can
CC also be used for in vitro production of biologically active proteins.
XX
SQ Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other:

Query Match 14.1%; Score 1148.2; DB 19; Length 14455;
Best Local Similarity 95.5%; Pred. No. 1.6e-162;
Matches 1196; Conservative 0; Mismatches 48; Indels 9; Gaps 1;

QY 755 gggcgccgctgagcagatgcatatagaggccattctatagtgacctaaatgctag 814
    |||||||
Db 11236 GCGGCCGCTCGAGCATGCATGTAGAGGCCCTATTCTATAGTGTACCAATATCTGTAG 11177
QY 815 ctgcctatagcctgacgtgacctgacctatagtcagcagcaatcggtgtttgcccctcc 874
    |||||||
Db 11176 CTGCTGTATCAGCCTCGACTGTGCTTCTTACTGTGCAACATCTGTGTGTCCTCCCTCC 11117
QY 875 ccgtgacctcttgacctggaaggtgccaactcccaactgctcttccataaaatgag 934
    |||||||
Db 11116 CCGAGCCTCTTCTTACCCCTGGAAGGTGCACCTCCACTGCTCTTCCATTAATTAATGAG 11057
QY 935 aaattgcacgtcattgtctgtagtgatcattctatctgagggtggtgggtggcaag 994
    |||||||
Db 11056 AAATTGCATCCATTGCTGTAGTGTGCTCATTTCTTGTGGGGGTGGGGTGGGAGG 10997
QY 995 acagcaaggaggagattgggaagaacataagcagcagctgctggagatgagcgtgagccta 1054
    |||||||
Db 10996 ACAGCAAGGGGAGGATTTGGGAACATAGCAGCATGCTGGGGAATCGGTGGGCTCTA 10937
QY 1055 tggcttctgaggcggaagaacacagctgcatlaatgaatcgccaacgctggggagagag 1114
    |||||||
Db 10936 TGGCTTCTGAGGCGGTA-----TGGGCTGTGAATATACCGACAGATGCTGAAGAG 10886
QY 1115 cgtttgcgtattggtgaccttccgcttctgctgctacgtacgtcgtgctggtcgt 1174
    |||||||
Db 10885 AAAATACCGATCAGGGCGCTCTCCGCTTCCCTCCTCATCTACATCGCTGCGCTGCGCT 10826
QY 1175 tcggtctgagcagcggtatagctcaactcaaaagcggtataacgttatccagatc 1234
    |||||||
Db 10825 TCGGCTGGGGGAGCGGATATAGCTCATCAAAAGCGGTATATACGGTTATCATCAGAAATC 10766
QY 1235 aggggatacgcaggaagaacatgtgagcaaaagcgacgcaaaagcgagaaacgttaa 1294
    |||||||
Db 10765 AGGGGATATACCGAGGAAGAATGTGAGCAAAAGCGCAGCAAAAGCGCAAGAACCGTAA 10706
QY 1295 aaagcgcgctgtgctggcgttttccataggctccgccccctgacgagcatcaaaaa 1354
    |||||||
Db 10705 AAAGGCGCGTGTGCGGCTTTTCCATATAGGCTCCGCGCCCTGACGAGCATCAAAAA 10646
QY 1355 tcgagctcaagtagaggtggtggaacccgacagagactataaagataccaggggtttcc 1414
    |||||||
```

D	b	10645	TCGACGGCTCAAGTCAAGAGGTGGCGAAACC	C	GCAACCCAGCAGACTATTAAGAATTCCAGGCCTTTCC	10586
O	y	1415	cgcctgaagctccctcgtgcgctcctcgttccagaacc	t	gcgcgttacggaatacgttc	1474
D	b	10585	CCCTTGGAAGCTCCCTCGTAGCGGTCTCTCGTTGCCA	C	CGCTTACCGGATTA	10526
O	y	1475	cgccttccctcccttgaggaaacgtagcgcttctcaat	g	ttcactcagctglaagtacag	1534
D	b	10525	CGCCTTCTCCCTTGGGGAAGCGGTGGCGCTTTCTCAT	A	GTACGCTCACGCTAGATATCTCAG	10466
O	y	1535	ttcagtgtaagtcgcttgctccaagctggctgtgtgtc	a	agaaaccccccggttcagccga	1594
D	b	10465	TTTCGGTGTAGGTCGTTCCCTCCAACCTGGGCTGTGTG	A	CGAACGCCCCGTTCCACCCGA	10406
O	y	1595	ccgcgcgcacctatccggttaactacacgcgtcttgagtc	a	cccgcggtgaagacacgactatc	1654
D	b	10405	CCGCTGCGCCCTTATCCGGTAACATACTGCTTGAGTCA	A	CCCGGTAAACACAGACTTATC	10346
O	y	1655	gccacttgcacaagcacgacttggtaaagaattagcaga	g	agcaggtlactgtagcgtgtac	1714
D	b	10345	GCCACTGGCAGAGCGCACTGGTGAACAGATTACGAC	A	CGAGCGAGTGTATGAGGGGTCTAC	10286
O	y	1715	agaattctttaagtagtggtgcctaactacgcgtacac	t	ctagaagacagtaattgtatctg	1774
D	b	10285	AGAGTTCCTTGAAGTGTGTGGCTTAACCTACGCGCTAC	A	CTGTGAAGACAGATATTGTGATCTG	10226
O	y	1775	cgcctctgttgaagccagttactcttcoggaaaagatt	g	tgtatgcttcttatccoggaaca	1834
D	b	10225	CGCTTGCTGAAGCCAGTTACCTTCGGAAAAAGATTG	T	GTAGCTTATCCGCGAACA	10166
O	y	1835	aaccacgcgttgaacggttgttlttctgttccaagcag	a	gatctaacgcagaanaaaa	1894
D	b	10165	AACCAACCGCTGTACCGGTGCTTTTTTTTGTTCGAC	A	GCAGCAGAGATTACGCCAGAAAAA	10106
O	y	1895	aggatctcaagaagaatcctttgatccttcttcaacgg	g	gtctgaagctcagtgaacyaaaa	1954
D	b	10105	AGGATCTCAAAAGATCTTGTGATCTTCTTCTAACG	G	GGTCTGACGCTCAGTGAACGAAAA	10046
O	y	1955	ctcacgttaagggatttgtgtcatagagcgatatact	a	attgaatgattataga	2007
D	b	10045	CTCACGTTTAAGGATTTGTGTATGATAGATTTATCA	A	AAAAGATCTTCACTTACG	9993
RESULT 7						
AAAS9050/c						
ID	AAAS9050 standard; DNA; 14455 BP.					
XX						
AC	AAAS9050;					
XX						
DT	07-NOV-2000 (first entry)					
XX						
DE	Nucleotide sequence of the EI/fiber-expressing plasmid pEI/fiber.					
XX						
KW	Adenovirus; tripartite leader; adenovirus vector particle; gene delivery					
KM	El gene; fiber gene; ss.					
OS	Synthetic.					
OS	Adenovirus type 5.					
XX						
PN	MO200042208-A1.					
XX						
PD	20-JUL-2000.					
XX						
PF	14-JAN-2000; 2000WO-EP00265.					
XX						
PR	14-JAN-1999; 99US-0115920.					
XX						
PA	(NOVS) NOVARTIS AG.					
FA	(NOVS) NOVARTIS-ERFINDUNGEN VERN GES MBH.					
XX	(SCRI) SCRIPPS RES INST.					
I1	Nemerow GR, Von Seggern DJ, Hallenbeck PL, Stevenson SC; Skripchenko Y;					

XX	WP1: 2000-476068/41.
DR	
XX	New nucleic acid comprising an adenovirus tripartite leader nucleotide
PT	for producing high-capacity and targeted vectors for adenovirus-based
PT	gene therapy -
XX	
PS	Example 1: Page 160-164; 212pp; English.
XX	
CC	The specification describes a nucleic acid molecule comprising an
CC	adenovirus (AV) tripartite leader (TPL) nucleotide with a sequence
CC	comprising two different TPL exons or three same or different TPL
CC	exons. The nucleic acid is used to produce an adenovirus vector particle
CC	deliver an exogenous gene to a target cell, pseudotype recombinant viral
CC	vectors, target an adenovirus vector to a cell, produce a modified
CC	adenovirus, deliver a heterologous gene to an animal and produce a
CC	gutless adenoviral vector particle. The present sequence represents
CC	a El/fiber-gene expressing plasmid, which was used for complementation of
CC	El/fiber-gene deleted adenoviruses.
XX	
Q0	Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other;

Query Match	14.1%	Score 1148.2	DB 21	Length 14455
Best Local Similarity	95.5%	Pred. No. 1.6e-162		
Matches 1196	Conservative	0	Mismatches 48	Indels 9
			Gaps	14
QY 755	gcggccgctcgaagcatgcatctatgagggcccatctctatagtggtacccaatagctagag	814		
Db 11236	GGGCGCGCTCGAGCATGCATCTAGAGGGCCCTATTCTATAGTGTACCTAAATGCTAGAG	11177		
QY 815	ctcgctgcatcagcctcgacatgctcctctctagttgcgaagcatalgtlgtltyccctccc	874		
Db 11176	CTCGGTATCATGACCCTCGACTGCTGCTTCAGTGTGCAGCAGCATCTGTGTTTCCCTCC	11117		
QY 875	ccgtgacctctcttgacccttggaaggtgcccacccacacgtcctctcccaataaatlgag	934		
Db 11116	CCGTCCCTTCTCTGACCCCTGGAAAGGTGCACCTCCACATGCTCTTCTCTATATAATGAG	11057		
QY 935	aaattgcatacgcatgctctgtagtagtgltcatctcatctctggygggtlggygtggcag	994		
Db 11096	AAATTGCATCCGATTTGCTTAGTAGTGCATCTATTCTGGGGGTGGGGTGGGGCAGG	10997		
QY 995	acagcaaggggggaagatttgggaagaacaatagcaggcatgcttggggatgcgttggctcta	1054		
Db 10996	ACAGCAAGGGGAGGATTTGGGAAGCAATAGCAGCATCTGGGGATCGGGTGGCTCTTA	10937		
QY 1055	tgagctctggaagcggaaagaaccagctgcatatgaatcggcaacgcgcggggagag	1114		
Db 10936	TGGCTTCTGAGGCGGTA-----TGGGTGTATAAATACCGCACAGATGGCTTAAGAG	10886		
QY 1115	cgglttgcatcttgaggcctctccgctctccgctctcgcgtcaactgcatcgcgttgcgtcgt	1174		
Db 10885	AAAATACCGCATTCAGGCGCTCTTCGCGTTCCTCGCTCACTGCATCGTGGCTCGGTCGT	10826		
QY 1175	tcgagctggggagagcggtatcatgctcactcaaaagcggtaaacggtatccaaagaatc	1234		
Db 10825	TCGCGTGGGGGAGCGGTATCACTCACTCAAAAGCGGTAATACGGTTATCCACAGAAATC	10766		
QY 1235	aggagataacgcaggaagaacaatcttgagcaaaagccagcaaaagcgccaggaacgctaa	1294		
Db 10765	AGGGATATACCGCAGGAAGAACAATGTGAGCAAAAGCCAGCAAAAGGCCAGGAACGTTAA	10706		
QY 1295	aaaggccgctgtctgctgacgtlcttccataagctccgcgccccctgagagcatcacaaaa	1354		
Db 10705	AAAGGCCCGCTGTTCGGCGGTTTTCATATAGCTCCGCCCCCTGTGACGATCACAAAAA	10646		
QY 1355	tcgagcgtccaaagtgaagtgtggcaaacccgaacagagcatataaagaatccaaagcgttcc	1414		
Db 10645	TCGACGCTCAAGTCAAGAGGTGCGAANAACCGACAGAGCATATAAATATACAGGCGTTTCC	10586		
QY 1415	cccttgaagctccctctgctgctcctctctgttccgagcccttgcgcgtctacggataactgtc	1474		

Db 10585 CCCTGGAAAGCTCCCTGCGCCTCTCTGTTCCAGACCTGCCGCTTACCGGATACCTGTC 10526
Qy 1475 cgccttcctccctcgggaagcgtgagcgtcttcctcaatcgtacagcgtgtagtalcag 1534
Db 10525 CGCCTTCTCCCTTGGGAAGGCTGGCGCTTCTCATAGCTCAGCTGAGGTATCTCAG 10466
Qy 1535 ttccggtgtagtgcgttcgtcctcaagcgtggtgctgtgtgacagaaaccccggtcacccga 1594
Db 10465 TTCGGTGTAGGTCGTTCCCTCCAAAGCTGGGCTGTGTGACAGAAACCCCGCTTACCCCGA 10406
Qy 1595 cgcgcgcgcctatccggttaactatcgtcttgagtcacaccccggtgaagacacgacttacc 1654
Db 10405 CGCGTGGCCCTTATCCGCTACTATCGTCTTGAGTCCAAACCCGGTAAGACACGACTTATC 10346
Qy 1655 gccactggaagcagcagcactggttaacagagtagcagagcagagtagttagagcggtgctac 1714
Db 10345 GCCACTGGCAGCAGCCACTGGTAAACAGATTAGCAGACGAGATATGTAGGCGGTGCTAC 10286
Qy 1715 agagtcttgaagtggtggtcctaaactacagcgtacacatagaaggaagatattgtgactctg 1774
Db 10285 AGAGTCTTGAAGTGTGGCTTAACACTAGCGCTACACTAGAAAGACAGATTTGTATCTGTG 10226
Qy 1775 cgcctcgtgaagcagcgttacctcgcgaagaaagattgtagctcttgatccgcgaaca 1834
Db 10225 CGCTGTGCTGAAGCCAGTTACTCTTGGAAAAAGACTTGTGACTCTTGTATCCGCCAACA 10166
Qy 1835 aaccacgcgtgtagcagtggttcttctgttgcagaagcagatlaacgcgcagaaaaa 1894
Db 10165 AACCCACCCCTGTACGGTGTGTTTTTTTGTGGCAAGCAGCAGATTTAGCGCGAGAAAAA 10106
Qy 1895 aggaactaagaagaatccttgcattcttctacagcggtcgtcagcgtcagtggaacgaaaa 1954
Db 10105 AGGATCTTAAGAAAGATCCTTGTGATCTTTTCTAGGGGCTGTGACGCTCAGTGAACGAAAA 10046
Qy 1955 ctcaagttaaaggaatttgatcatagcggatcacatattgaatgattaga 2007
Db 10045 CTCACGTTAAGAGGATTTTGTGTCATGAGATTAACAAAAAGATCTTCACCTAGA 9993
RESULT 8
ABA94259/c
ID ABA94259 standard; DNA; 14455 BP.
XX
AC ABA94259;
XX
DT 13-MAR-2002 (first entry)
XX
DE Nucleotide sequence of expression plasmid pE1/Fiber.
XX
KW Adenovirus; Inverter terminal repeat sequence; ITRs; ocular disease;
KW fiber protein; photoreceptor; rhodopsin; starardt disease gene; STDG1;
KW ophthalmological; antiinflammatory; antidiabetic; cytoskeletal;
XX
OS Synthetic.
XX
PN WO200183729-A2.
XX
PD 08-NOV-2001.
XX
PE 30-APR-2001; 2001WO-EP04863.
XX
PR 01-MAY-2000; 2000US-0562934.
XX
PA (NOVS) NOVARTIS AG.
PA (SCRI) SCRIPS RES INST.
PA (NEME/) NEMEROW G R.
PA (VSEG/) VON SEGGERN D J.
PA (FRIE/) FRIEDLANDER M.
PI Nemerow GR, Von Seggern DJ, Friedlander M;
XX
DR WPI: 2002-082846/11.

XX
PT Polynucleotide for making vectors, useful for treating ocular diseases,
PT e.g., retinitis pigmentosa, comprises adenovirus inverter terminal
PT repeat sequences, packaging signal and photoreceptor-specific promoter
PT
XX
PS Example 1: Page 114-118; 149pp; English.
XX
CC The invention provides an isolated polynucleotide comprising adenovirus
CC (AV) inverter terminal repeat sequences (ITRS), AV packaging signal
CC operatively linked to ITRS and a photoreceptor-specific promoter. A
CC recombinant AV vector (AAV) comprising the polynucleotide is useful for
CC targeted delivery of a gene product to the eye (especially to the
CC vitreous cavity), for treating an ocular disease, e.g., retinal
CC degenerative disease, retinitis pigmentosa, Starardt's disease, diabetic
CC retinopathies, retinal vascularizations, and retinoblastoma, of a mammal
CC preferably human. The AAV comprises a fiber protein that specifically or
CC selectively binds to receptors that are expressed on cells (preferably
CC photoreceptors in the eye). Preferably, the recombinant virus comprise a
CC fiber protein from an adenovirus type B subgroup or is a chimeric protein
CC containing a portion of the N-terminus of an adenovirus type 2 or type 5
CC penton, and the therapeutic product is a trophic factor, an anti-
CC apoptotic factor, gene encoding a rhodopsin protein, a wild-type
CC starardt disease gene (STDG1), an anti-cancer agent and a protein that
CC regulates expression of a photoreceptor specific gene product. The viral
CC nucleic acid of AAV comprises ITRS and packaging signal derived from AAV
CC subgroup B or C, especially an AV type 2 or type 5. AAV is also useful
CC for targeted gene therapy, where the vector comprises an AV type 37 fiber
CC protein or its portion, and selectively transduces photoreceptors and
CC delivers a gene product encoded by AAV. The present sequence represents
CC an expression plasmid pE1/Fiber containing the adenovirus EI and fiber
CC gene.
SQ Sequence 14455 BP; 3698 A; 3271 C; 3565 G; 3921 T; 0 other;
XX
Query Match 14.1%; Score 1148.2; DB 24; Length 14455;
Best Local Similarity 95.5%; Pred. No. 1.6e-162;
Matches 1196; Conservative 0; Mismatches 48; Indels 9; Gaps 1;
Qy 755 ggcgcgcgtcagcatgatactagaagggccctatctatagtgtaacataatgtagag 814
Db 11236 GCGCGCGCTCGAGCATGATGATGAGGCGCCATTCATAGTGCATTAATGCTAGAG 11177
Qy 815 ctgcgtatcagcctcgcactgtgcttctagtgtagtcagcaatcgtgtttgcccctcc 874
Db 11176 CTCGCTGATCAGCCCTCGACCTGCTGCTTCTAGTTGACACCATCTGTTTGGCCCTCC 11117
Qy 875 cgcgtccttccttgcacctggaagtgccactcccaatgcttcctccaataaataagag 934
Db 11116 CCGTCCCTTCTTGACCTCGGAAGGTGCACCTCCACTGCTCTTCCCTAATAAATGAGG 11057
Qy 935 aaatgcacatcgtctgtagtgtagtgcattctatctcttgagggtggtggtgagcag 994
Db 11056 AAATGCATCGATTTGTGTAGTAGTGTCTCATTTCTGAGGGGTGGGGGCGGCAAG 10997
Qy 995 aacgaagggggaggaattgggaagacaalagcaagcagtcgtggtggtggtggtggtggt 1054
Db 10996 ACAGCAAGGGGAGGATTTGGGAAGCAATPACAGGCAATGCGGATCGGTGGCTCTTA 10937
Qy 1055 tggcttcgtaggctggaagaaacagctgcatlaatgaatcgcccaacgcggtgggaagag 1114
Db 10936 TGGCTTCTGAGGCGGTA-----TCCGGTGTGAATPACCGCAACATCGTAAAGAG 10886
Qy 1115 cgtttgctgtagtggtggtccttcgcgtcttcgtctgaactgaactgcgtgcgtgctgt 1174
Db 10885 AAATACCGCATCAGGCCCTTTCGCTTCTCGTCTCATCTGACTGCTGCGCTCGGTGCT 10826
Qy 1175 tcgcgtcgcgcagcagcgtatcaactcaaaagcgtgtaataagttatccagaatc 1234
Db 10825 TCGGCTGGGAGGAGGATATCAGCTCAAGAGCGGTAAATAGGTTATCAGCAGATTC 10766
Qy 1235 aggggataacgcaggaagaacatgtgagcaaaagggccagcaaaagccaggaacgttaa 1294

```

Db 10765 AGGGAGTAACGACGAGAAAGAAATGTAGAGCAAAAGGCCAGCAAAAGGCCAGAACCGTAA 10706
|
|
|
Qy 1295 aaagcgcggttctgtcgcttccatagagctccgccccctgaagacatcaaaaa 1354
|
|
|
Db 10705 AAAGCGCGGTTGCTGGCGTTTTCATAGGCTCCGCCCTGACGAGCATCAAAAA 10646
|
|
|
Qy 1355 tcgagcgtcaagtcagaggtggtggaacccgacagactataaagataccagaggttcc 1414
|
|
|
Db 10645 TCGAGCTCAAGTCAGAGGTGGCAAAACCGACAGACTATAAAGATPACGAGCGTTTCC 10586
|
|
|
Qy 1415 cccgtgaagctccctctgtcgctcctgttccagaccctgcgcttaaccgataccctgtc 1474
|
|
|
Db 10585 CCTGGAAGCTCCCTCGTGGCGCTCTCTGTTCACACCCCTGCGCTTACCGGATACCTGTC 10526
|
|
|
Qy 1475 cgccttctccctcgtggaagcgtgtgcgcttctcattgtctacgctgtatgtctcag 1534
|
|
|
Db 10525 CGCCTTCTCCCTTCCGGGAAGCGTGGCGCTTTCATAGCTCACGCTGTAGTATCTCAG 10466
|
|
|
Qy 1535 ttgcgtgtagtctgtcgtcccaagctgggtctgtgtacgaaccccgcttgcgcga 1594
|
|
|
Db 10465 TTGCGGTAGGTCCTTGGCTTCCAGCTGGGCTGTGTGCACAAACCCCGTTACGCCGA 10406
|
|
|
Qy 1595 ccgctgcgcttaccggttaactcgttctgaagtcacacccggttaagacacgactatc 1654
|
|
|
Db 10405 CCGCTGCGCTTATCCGCTACCTATCGTTCGTGAGTCCAAACCCGTTAAGACAGACTTATC 10346
|
|
|
Qy 1655 gccactgagcagcagcactgtgaagatagcagaagcgaagtatgtagcggtgtac 1714
|
|
|
Db 10345 GCCACTGTGGACACCACTGTAAACAGATTAGCAGACGAGGATNTAGCGGCTGTCTAC 10286
|
|
|
Qy 1715 agacttctgaagtgtgtgccttaactacggtctacactagaagacagtatgtgtatctg 1774
|
|
|
Db 10285 AGACTTCTTAAGAGGTGGCTTAACCTACGGCTACACTAGAAAGACATATTGGTATCTG 10226
|
|
|
Qy 1775 cgcctgtcgaagcagcttacctcggaaaaagagtgtagctctgtacctcggcaaca 1834
|
|
|
Db 10225 CGCCTCTCTAAGACCACTTACCTTCGGAATAAGATTGTTAGTCTTGATCCGCAACA 10166
|
|
|
Qy 1835 aaccacgcgtgtgtagcggtgttctgttctgaacgacgatacgcgcaaaaaa 1894
|
|
|
Db 10165 AACCAACCGCTGTAGCGGTGTTTTTTTGTTCAGACAGCAATWTAACGGCAGAAAAA 10106
|
|
|
Qy 1895 aggaatccaaagaatcccttgcattcttctacggtgtcgaacgtcagtcgaacgaaa 1954
|
|
|
Db 10105 AGGATCTCAAGAGATCTTTGATCTTTCTACGGGCTCTGAGCCTCAGTGGAAAGAAA 10046
|
|
|
Qy 1955 ctcaagcttaaggaattgtgcatgagcgatatacatttgaaatgtattaga 2007
|
|
|
Db 10045 CTCACGTTAAGGATTTTGGTCAATGATTAATAAAAAAGATCTTCACCTAGA 9993
|
|
|

```

RESULT 9

```

AAAS3872
ID AAAS3872 standard; DNA: 5377 BP.
AC AAAS3872:
XX
XX 03-JAN-2001 (first entry)
DE Expression vector PRIG14.
XX
XX Vector; endogenous gene; activation; over-expression;
KM erythropoietin; growth hormone; drug discovery;
KM granulocyte colony stimulating factor; ds.
XX
XX Synthetic.
XX
XX WO200049162-A2.
XX
XX 24-AUG-2000.
XX
XX 22-FEB-2000; 2000MO-US04429.
PF

```

```

XX
PR 19-FEB-1999; 99US-0253022.
PR 08-MAR-1999; 99US-0263814.
PR 26-MAR-1999; 99US-0276820.
XX
PA (ATHE-) ATHERSYS INC.
XX
PI Harrington JF, Sherf B, Rundlett S;
XX
XX WPI; 2000-549276/50.
XX
PT Non-targeted activation of endogenous genes, e.g. for the production of
PT erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT proteins and for drug discovery
XX
PS Disclosure; Fig 29; 240pp; English.
XX
CC New methods, vectors and cells are described for non-targeted
CC activation and over-expression of endogenous genes. The vector
CC constructs comprise transcriptional regulatory sequences (TRS) and
CC unpaired splice donor sequences (USDS), preferably the vectors
CC comprise (in sequential order) a TRS, an USDS, a rare cutting
CC restriction site (RCRS) and a linearization site (LS) with a second
CC TRS linked to a selectable marker (SM) lacking a polyadenylation
CC signal. The methods, vectors and cells comprising the vectors may
CC be used for the non-targeted activation and over-expression of
CC endogenous genes, e.g. for the production of proteins (including
CC erythropoietin, growth hormone or granulocyte-colony stimulating
CC factor) and drug discovery. The advantage of these methods are that
CC endogenous genes including those associated with human disease and
CC development, may be activated and isolated without prior knowledge
CC of the sequence structure, function or expression profile of
XX the genes being known.
XX
SQ Sequence 5377 BP; 1248 A; 1465 C; 1437 G; 1225 T; 2 other:

Query Match 14.0%; Score 1140; DB 21; Length 5377;
Best Local Similarity 74.3%; Pred. No. 3e-161;
Matches 1684; Conservative 0; Mismatches 300; Indels 284; Gaps 8;

Qy 23 gttgacattgatttactagctagtattataatagtaatcaattacgggtcatgtatcata 82
|
|
|
Db 2058 gttgacattgatttactagctagtattataatagtaatcaattacgggtcatgtatcata 2117
|
|
|
Qy 83 gcccatatagggttcgggttacataacttaaggaatggtccgctggcgaccgc 142
|
|
|
Db 2118 gcccatatagggttcgggttacataacttaaggaatggtccgctggcgaccgc 2177
|
|
|
Qy 143 ccaagaccccgcccatatgacgtcaataatgacgtatgtcccatgtaacgcaatag 202
|
|
|
Db 2178 ccaagaccccgcccatatgacgtcaataatgacgtatgtcccatgtaacgcaatag 2237
|
|
|
Qy 203 ggaattccattgacgtcaatgggtgactatttaacggtaaactggccacttgagctac 262
|
|
|
Db 2238 ggaattccattgacgtcaatgggtgactatttaacggtaaactggccacttgagctac 2297
|
|
|
Qy 263 atcaagtgatcatatgccaagtacgccccctattgacgtcaatgaaagaaatggtccg 322
|
|
|
Db 2298 atcaagtgatcatatgccaagtacgccccctattgacgtcaatgaaagaaatggtccg 2357
|
|
|
Qy 323 cctgcatatgcccagtaacgtacacttaaggaacttccacttggcagtaacacag 382
|
|
|
Db 2358 cctgcatatgcccagtaacgtacacttaaggaacttccacttggcagtaacacag 2417
|
|
|
Qy 383 tattagtcacgtcattacacatggtgtagtgcgttttggcagtaacatggtggcgtgat 442
|
|
|
Db 2418 tattagtcacgtcattacacatggtgtagtgcgttttggcagtaacatggtggcgtgat 2477
|
|
|
Qy 443 agcggttgactcaaggggatttccaaagtctccacccattgacgtcaatggtggattgtc 502
|
|
|
Db 2478 agcggttgactcaaggggatttccaaagtctccacccattgacgtcaatggtggattgtc 2536
|
|
|

```

[illegible]

Dd	3617	 ccgcgtctgctggcgttcttcctaagcgctcccgccttgacgaagatcaacaataactgac	3676
Oy	1360	gctcaagaccagaaggctggcgaacccgcgaagacctaaagaataccaagcgctttccccctg	1419
Dd	3677	gctcaagtcagaaggctggcgaacccgcgaagacctataaagaataccaagcgctttccccctg	3736
Oy	1420	gaagctccctctgctgctctctgcttcgcgaacctgcgcgttaacgatattgctgcgcct	1479
Dd	3737	gaagctccctctgctgctctctgcttcgcgaacctgcgcgttaacgatattgctgcgcct	3796
Oy	1480	tctccctctgggaagcgttgccgtcttctcaatgctccagctgttagtatctcaagtctgg	1539
Dd	3797	tctccctctgggaagcgttgccgtcttctcaatgctccagctgttagtatctcaagtctgg	3856
Oy	1540	tgtaatgctgtctgcttccaaaagcttgagctgtgtgacgaaccccccgcttaagccgaacgct	1599
Dd	3857	tgtaatgctgtctgcttccaaaagcttgagctgtgtgacgaaccccccgcttaagccgaacgct	3916
Oy	1600	ggcgcttatccggtgaactatcgtctgtgagtcacaaccggttaagacagacttatctgcac	1659
Dd	3917	ggcgcttatccggtgaactatcgtctgtgagtcacaaccggttaagacagacttatctgcac	3976
Oy	1660	tggcacgacgacacactggttaaacaggtatgacgaagcgaagtatgaagcsgtgtcatagagt	1719
Dd	3977	tggcacgacgacacactggttaaacaggtatgacgaagcgaagtatgaagcsgtgtcatagagt	4036
Oy	1720	tcttgaaagtgtgtggcctaactacgcttaacatagaagaagactattggtatctgcgctc	1779
Dd	4037	tcttgaaagtgtgtggcctaactacgcttaacatagaagaagactattggtatctgcgctc	4096
Oy	1780	tgtctaaagccagcttaccttcggaaaaaagttgtagctccttgatcccgcaacaacaaca	1839
Dd	4097	tgtctaaagccagcttaccttcggaaaaaagttgtagctccttgatcccgcaacaacaaca	4156
Oy	1840	cgcgctgtgaagcsgtgtgttttttgtgttgcaagcagcagatacgcgcgaagaaaaaagat	1899
Dd	4157	cgcgctgtgaagcsgtgtgttttttgtgttgcaagcagcagatacgcgcgaagaaaaaagat	4216
Oy	1900	ctcaagaagaatcccttatactttctacaggggtctgacgctacagtcgaagcaaaaaacttac	1959
Dd	4217	ctcaagaagaatcccttatactttctacaggggtctgacgctacagtcgaagcaaaaaacttac	4276
Oy	1960	gttaagggatttggltcatbagcgaatactttaatgatatltaga 2007 	
Dd	4277	gttaagggatttggltcatbagcgaatactttaatgatatltaga 4324 	
RESULT 10			
AAAS3870	ID	AAAS3870 standard; DNM; 4645 BP.	
XX AC	AAAS3870;		
XX XX	03-JAN-2001 (first entry)		
DE	Expression vector pRIC21b.		
XX			
KM	Vector: endogenous gene; activation; over-expression;		
KW	erythropoietin; growth hormone; drug discovery;		
KX	granulocyte colony stimulating factor; ds.		
OS	Synthetic.		
PN	WO200049162-A2.		
PD	24-AUG-2000.		
PF	22-FEB-2000; 2000MO-USO4429.		
PR	19-FEB-1999; 99US-0253022.		
PR	08-MAR-1999; 99US-026314.		
PR	26-MAR-1999; 99US-0276820.		

RESULT 12
AAS18619
ID AAS18619 standard; DNA; 5446 BP.
XX
XX
AC AAS18619;
XX
DT 26-FEB-2002 (first entry)
XX
XX
DE Renilla luciferase vector pCDNA3.
XX
KW Renilla luciferase; sea pansy; cell proliferation disorder;
KW immune disorder; hypogammaglobulinemia; hematologic condition;
KW anaemia; neoplasm; cancer; human immunodeficiency virus; HIV;
KW tissue white cell infiltrative disorder; organ failure;
KW myotrophic condition; gonadal failure; bone disorder; muscle disorder;
KW osteoporosis; endocrine condition; vascular disorder; atherogenesis;
KW pCDNA3; ds.
XX
XX
OS Synthetic.
XX
PN W0200181614-A2.
XX
PD 01-NOV-2001.
XX
PE 25-APR-2001; 2001WO-US13512.
XX
PR 25-APR-2000; 2000US-0559874.
XX
PR 02-JUN-2000; 2000US-0586339.
XX
PA (CHEM-) CHEMICON INT.
XX
PI Leng J;
DR WPI; 2002-041420/05.
XX
PT Determining cell proliferation for monitoring treatment of a subject,
PT comprises obtaining light emission data from cell containing Renilla
PT luciferase for specific time, and detecting a change in the data
XX
PS Example 1; Fig 1A-B; 52pp; English.
CC
CC The invention describes a novel method for measuring proliferation of a
CC cell or population of cells. The method comprises obtaining light
CC emission data from a cell containing a Renilla luciferase over a period
CC of time, cell proliferation of a cell or a population of cells can be
CC measured by a change in light emission data indicating proliferation. A

Query Match	13.3%	Score 1082;	DB 24;	Length 5446;
Best Local Similarity	100.0%;	Pred. No. 1,3e-152;		
Matches 1082;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Oy	1	gatgacgggcacagatagcagcttgacattgattatgtactatgaatgaataa	60	
Db	210	gatgtacgggcccagagataacgcgttgacattgattatgtactatgaatgaataa	269	
Oy	61	attaaggggtcattatgattcatagccacataatggagttccgcgtttacataacttcg	120	
Db	270	attacgggtcattatgattcatagccacataatggagttccgcgtttacataacttcg	329	
Oy	121	aatgcccgcctgctgtctgacccgcccaagaccccccgcatttgaagctcaataatgac	180	
Db	330	aatggccgcctgtgtctgacccgcccaagaccccccgcatttgaagctcaataatgac	389	
Oy	181	gttcccaatgataagccacataaggagcttccattgacgtcaatgggtgagactattg	240	
Db	380	gttcccaatgataagccacataaggagcttccattgacgtcaatgggtgagactattg	449	
Oy	241	taaatgtcccaatttggcagtatcatcaatgattatcatatgccaagtlacgccccctat	300	
Db	450	taaatgtcccaatttggcagtatcatcaatgattatcatatgccaagtlacgccccctat	509	
Oy	301	gtcaatgacggtataatggccgcgcctgtgcattatgtcccaatgatagtaccta	360	
Db	510	gtcaatgacggtataatggccgcgcctgtgcattatgtcccaatgatagtaccta	569	
Oy	361	cctacttggcagttacatctacgtattagttacgtctataccatgaggtacggtttg	420	
Db	570	cctacttggcagttacatctacgtattagttacgtctataccatgaggtacggtttg	629	
Oy	421	cagttacatcaatltgggcgtgtgaatagcgttttgaactacacgggatttccaagtctcc	480	
Db	630	cagttacatcaatltgggcgtgtgaatagcgttttgaactacacgggatttccaagtctcc	689	
Oy	481	attgacgtcaatgtggaggtgttttttggcaccaaaatacaaggagacttccaaatgtc	540	
Db	690	attgacgtcaatgtggaggtgttttttggcaccaaaatacaaggagacttccaaatgtc	749	
Oy	541	aacaaactcggcccaattgacgcaaatgggcgttaggcgtgtgaacggtgggggtctata	600	
Db	750	aacaaactcggcccaattgacgcaaatgggcgttaggcgtgtgaacggtgggggtctata	809	
Oy	601	agcgaagctctctgtgctaactagaaacccactgtcttactgtcttaccgaatgaatac	660	

D8 810 agcagagctctctgctaactagagaccacactgctactgtgtctatctcgaataatcgcg 865
QY 661 actactactagagagaccacaaagtgtgtaccagagctcgatccactagtaacgcccga 720
D8 870 actactactagagagaccacaaagtgtgtaccagagctcgatccactagtaacgcccga 929
QY 721 gtgtgtgtggaattctgtcagatatacattacacactgtggtcgctcgagacatgtcctag 780
D8 930 gtgtgtgtggaattctgtcagatatacattacacactgtggtcgctcgagacatgtcctag 989
QY 781 ggcacctatctatagtgtcacactaatgtcagagctcgatccactagctgactgtgct 840
D8 990 ggcacctatctatagtgtcacactaatgtcagagctcgatccactagctgactgtgct 1049
QY 841 tctagtgtcagcagcactgt 900
D8 1050 tctagtgtcagcagcactgt 1109
QY 901 ggcacctcactgtcctctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 960
D8 1110 ggcacctcactgtcctctcctcctcctcctcctcctcctcctcctcctcctcctcctcct 1169
QY 961 tgtcattctatctctgt 1020
D8 1170 tgtcattctatctctgt 1229
QY 1021 aatagcagcagctgt 1080
D8 1230 aatagcagcagctgt 1289
QY 1081 tg 1082
D8 1290 tg 1291

RESULT 13

AAA53871
ID AAA53871 standard; DNA: 5245 BP.

AC AAA53871;

DT 03-JAN-2001 (first entry)

DE Expression vector pRIG22b.

XX Vector; endogenous gene; activation; over-expression;

KW erythropoietin; growth hormone; drug discovery;

KW granulocyte colony stimulating factor; ds.

OS Synthetic.

PN WO200049162-A2.

PD 24-AUG-2000.

PF 22-FEB-2000; 2000WO-US04429.

PR 19-FEB-1999; 98US-0253022.

PR 08-MAR-1999; 99US-0263814.

PR 26-MAR-1999; 99US-0276820.

PA (ATHE-) ATHERSYS INC.

PI Harrington JJ, Sherif B, Rundlett S;

DR WPI; 2000-549276/50.

XX Non-targeted activation of endogenous genes, e.g. for the production of
PT erythropoietin, growth hormone or granulocyte-colony stimulating factor
PT proteins and for drug discovery

XX Disclosure; Fig 16; 240pp; English.

CC New methods, vectors and cells are described for non-targeted
CC activation and over-expression of endogenous genes. The vector
CC constructs comprise transcriptional regulatory sequences (TRS) and
CC unpaired splice donor sequences (USDS), preferably the vectors
CC comprise (in sequential order) a TRS, an USDS, a rare cutting
CC restriction site (RCRS) and a linearization site (LS) with a second
CC TRS linked to a selectable marker (SM) lacking a polyadenylation
CC signal. The methods, vectors and cells comprising the vectors may
CC be used for the non-targeted activation and over-expression of
CC endogenous genes, e.g. for the production of proteins (including
CC erythropoietin, growth hormone or granulocyte-colony stimulating
CC factor) and drug discovery. The advantage of these methods are that
CC endogenous genes including those associated with human disease and
CC development, may be activated and isolated without prior knowledge
CC of the sequence structure, function or expression profile of
CC the genes being known.
XX
SQ Sequence 5245 BP; 1287 A; 1337 C; 1367 G; 1253 T; 1 other;

Query Match 13.28; Score 1071.6; DB 21; Length 5245;
Best Local Similarity 75.4%; Pred. No. 4.5e-151;
Matches 1503; Conservative 0; Mismatches 349; Indels 142; Gaps 8;

QY 23 gtgtacattgttttactagtagtattatagatataatcaggggtcattatata 82
D8 140 gtgtacattgttttactagtagtattatagatataatcaggggtcattatata 199
QY 83 gccatatatggaattccgcgttccataacttaacgttaacggccgcctgtgcagcgc 142
D8 200 gccatatatggaattccgcgttccataacttaacgttaacggccgcctgtgcagcgc 259
QY 143 ccaacgaccccccgcattgacgttcaataatgacgtatgttccatagtaacgcaatag 202
D8 260 ccaacgaccccccgcattgacgttcaataatgacgtatgttccatagtaacgcaatag 319
QY 203 ggaattccattgacgttcaatggtgtgactattagagtaacggccactgtgcagtac 262
D8 320 ggaattccattgacgttcaatggtgtgactattagagtaacggccactgtgcagtac 379
QY 263 atcaagtgatcatatgccaagtagcccccattgacgttcaatgacgttaacgttaacgccc 322
D8 380 atcaagtgatcatatgccaagtagcccccattgacgttcaatgacgttaacgttaacgccc 439
QY 323 cctgtcattatgcccagtagcatgactatgtgactttcctactgtgcagtaactacg 382
D8 440 cctgtcattatgcccagtagcatgactatgtgactttcctactgtgcagtaactacg 499
QY 383 tattagatcgtcattacccatgtgtatgc-ggtttggcagatcatcaatgggcgtgga 441
D8 500 tattagatcgtcattacccatgtgtatgc-ggtttggcagatcatcaatgggcgtgga 559
QY 442 tagcggtttgactacaggggaattcccaagttccacccattgacgttcaatgggaattg 501
D8 560 tagcggtttgactacaggggaattcccaagttccacccattgacgttcaatgggaattg 619
QY 502 ttgtggacccaataacagggactttccaaaatggtcgttaacact-----cgccc 553
D8 620 ttgtggacccaataacagggactttccaaaatggtcgttaacactttcgtacgcgcc 679
QY 554 cattgacgcaaatggcggtagcggtttagcggttggaggtcattataagcagagctcct 613
D8 680 cgttgaacgcaaatggcggtagcggtttagcggttggaggtcattataagcagagctcct 739
QY 614 ggtactactagagaaccactgtctactgtctactcgaataatcaactacactacataggg 673
D8 740 tagtgaacggtcagatcatcactagaagcttattgtggtcagttatccacagttta 791
QY 674 agaccacgaactgtgtacagagctcgtacactagtaacggcgccagtggtgtgaatt 733
D8 792 aattgtacagcagtcagtgcttctcgtacacacagctcgtacactaagctcgtcagtc 851
QY 734 ctgcagatatacattcaactgtgcggtcgtcagatgcatctagtaggccccttctat 793

```

Db      852  tcttaataccacccatggtacaggtgagtgactcgtactcaacttaagagagcgctatctggcc  911
Qy      794  aggttcaccataatgctcagagctcgcgtacagcctcgactgctgcttcaagtgtccagc  853
Db      912  agttaagcagctcgaagaagaagtttaagaagcgcaacaacagcgtcatatgagcccgag  971
Qy      854  catctgtgttctgcccctcccccgtgcttcccttgaccctggaaagggcgcaactcccaat  913
Db      972  ggcgagcccgatcttcccctcgtgtatgtcgcgcgatatagggccagcaacgcgacctg  1031
Qy      914  tcccttccctaataaaltgaggaattgcatcgcatgtctgtagtaggtgcatctatlc  973
Db      1032  t-----  1032
Qy      974  ttgggggtgtgggtgagcagagacgcaaggggagagattggagaaacaatagcagcagt  1033
Db      1033  ---gcgcgcgtgtatgtccgcgcagcatgctgcgcgttagagatccacagagcgggtg  1088
Qy      1034  ctggggatgcggtgaggtctatgtctctgagggcggaagaacagctgcatlaatgaat  1093
Db      1089  tggtcgcgaatgat-----cgcgtagtcgatagtg  1117
Qy      1094  cggccaaacgcgcgggagagagcggttgcgtatgtggcgctctccgcctccgcctcac  1153
Db      1118  gtcccaagtagcgaagcagcagcagcagctggtgcgcgcgaagcggtcgcgaagtgtcccg  1177
Qy      1154  tgcactgcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc  1213
Db      1178  agaacgggtgcgcatagaaattgcatcaacatatagcgtatgactctgtcgtcgtcgtc  1237
Qy      1214  aatacgttataccacagaatcaggggatagcgaagaagaacatgtgagcaaaaggcca  1273
Db      1238  agatctgtcgaag-----catggaagcaaaaggcca  1268
Qy      1274  gcaaaagccaggaacccgtaaaaaagcgcgctgtgcgttcttccataggtccgcgcc  1333
Db      1269  gcaaaagccaggaacccgtaaaaaagcgcgctgtgcgttcttccataggtccgcgcc  1328
Qy      1334  cccttgacgaacatcaaaaaaatacgaagctcaaatcagaagttgagcgaaccgcgaagact  1393
Db      1329  cccttgacgaacatcaaaaaaatacgaagctcaaatcagaagttgagcgaaccgcgaagact  1388
Qy      1394  ataagaatacagagcgcttcccccgtgaagctccctcgtcgtcgtcgtcgtcgtcgtcgtc  1453
Db      1389  ataagaatacagagcgcttcccccgtgaagctccctcgtcgtcgtcgtcgtcgtcgtcgtc  1448
Qy      1454  gccgcttaacggaataccctgtcgcgccttcccttcgcgggaagcgtgtgcgttcttccaat  1513
Db      1449  gccgcttaacggaataccctgtcgcgccttcccttcgcgggaagcgtgtgcgttcttccaat  1508
Qy      1514  ctacagctgtaggtatctcagttcgtgtgtatgtgtgtgtgtgtgtgtgtgtgtgtgtgt  1573
Db      1509  ctacagctgtaggtatctcagttcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt  1568
Qy      1574  cgaaccccccgttccagcccgacgcgtgcgcgtctatcccgtaactatcgtcttgagttccaa  1633
Db      1569  cgaaccccccgttccagcccgacgcgtgcgcgtctatcccgtaactatcgtcttgagttccaa  1627
Qy      1634  ccccggttaagaacagactatcgcaactgagcagcagccaactgtgtaacagattagcagagc  1693
Db      1628  ccccggttaagaacagactatcgcaactgagcagcagccaactgtgtaacagattagcagagc  1687
Qy      1694  gaaggtatgtagcgtgtcgtacagagttcttgaagttgttgccctaactagcgtacactag  1753
Db      1688  ga-gtatgtatagcgtgtcgtacagagttcttgaagttgttgccctaactagcgtacactag  1746
Qy      1754  aaggacagtaattgtgtatctgcgtctcgtcgtgaagccagttacaccttcggaaaaagattg  1813
Db      1747  aaggacacataattgtgtatctgcgtctcgtcgtgaagccagttacaccttcggaaaaagattg  1806
Qy      1814  tagctcttgatccggcaaaaacacgcgctgtgtaagcgtgtgtttttgttttgcaagca  1873

```

```

Db      1807  tagctcttgatccgcgcaaaaacacccgcgtgtgtaagcgtgtgtttttgttttgcaagca  1866
Qy      1874  gcagattacgcgcgaagaaaaaagatctcaagaagatcctttgattcttttctcgggtc  1933
Db      1867  gcagattacgcgcgaagaaaaaagatctcaagaagatcctttgattctttctcgggtc  1926
Qy      1934  tgaacgtcaattggaacgaacacacgttaaggattgtgtcatagcgcgaataatt  1993
Db      1927  tgaacgtcaattggaacgaacacacgttaaggattgtgtcatagcgcgaataatt  1986
Qy      1994  tgaatgbattaga 2007
Db      1987  gatcttcactaga 2000

RESULT 14
AAV40007
ID      AAV40007 standard; DNA: 4026 BP.
XX
XX
AC      AAV40007;
XX
DT      15-FEB-1999 (first entry)
XX
DE      Plasmid pCTMI.
XX
KW      E2F: transcription factor; human; retinoblastoma protein RB;
KW      bladder cancer; restenosis; angioplasty; diabetic retinopathy;
KW      thyroid hyperplasia; Grave's disease; psoriasis;
KW      benign prostatic hyper trophy; Li-Fraumeni syndrome;
KW      peripheral vascular disease; therapy; plasmid pCTMI; ss.
XX
OS      Chimeric - cytomegalovirus.
OS      Chimeric - mastadenovirus.
OS      Chimeric - bacteriophage T7.
OS      Chimeric - bacteriophage SP6.
OS      Chimeric - rhesus macaque polyoma virus.
OS      Chimeric - Bos taurus.
XX
FH      Key
FT      Location/Qualifiers
FT      promoter
FT      209..864
FT      /tag= a
FT      /note= "CMV promoter"
FT      misc_feature
FT      907..1074
FT      /tag= b
FT      /function= tripartite leader sequence
FT      intron
FT      1075..1253
FT      /tag= c
FT      /note= "hybrid SV40 late intron"
FT      promoter
FT      1305..1322
FT      /tag= d
FT      /note= "SP6 promoter"
FT      misc_feature
FT      1851..4026
FT      /tag= e
FT      /note= "pUC19 backbone R3 to AattII"
FT      CDS
FT      complement (3032..3890)
FT      /tag= f
FT      /note= "AMP-ORF"
XX
PN      WO9821228-A1.
XX
PD      22-MAY-1998.
XX
PF      13-NOV-1997; 97WO-US21821.
XX
PR      14-FEB-1997; 97US-0801092.
PR      15-NOV-1996; 96US-0751517.
XX
PA      (CANF-) CANJ1 INC.
XX
PI      Antelman D, Gregory RJ, Willis KN.
XX
XX      WPI; 1998-297858/26.
XX

```

PT New fusion polypeptide of, e.g. transcription factor - used to
PT treat, e.g. hyper-proliferative disease such as cancer and
PT restenosis

XX Example 1; Fig 6; 91pp; English.

CC This is the nucleotide sequence of pCTM1, a plasmid that was
CC constructed from pCTM (see AAV4006) by digesting pCTM with XhoI and
CC NotI and subcloning a 180 bp intron XhoI-NotI fragment from a
CC PCMV-beta-gal vector. Plasmid pCTM1 has been used as a vector for
CC the expression of fusion proteins of the invention that comprise
CC retinoblastoma protein (BP, see AAV62465) and E2F transcription
CC factor (see AAV62464). Such fusion proteins, particularly expressed
CC from gene therapy vectors, are used to treat hyperproliferative
CC conditions, specifically cancer (particularly of the bladder) or
CC restenosis. They are more effective in repressing transcription of
CC the E2F promoter than RB alone and cause cell-cycle arrest in a
CC variety of cells.

XX Sequence 4026 BP; 978 A; 1021 C; 982 G; 1045 T; 0 other;

Query Match 13.2%; Score 1070.2; DB 19; Length 4026;

Best Local Similarity 69.7%; Pred. No. 7.4e-151;

Matches 1919; Conservative 0; Mismatches 88; Indels 746; Gaps 5;

QY 1 gatgtacggccagatatacagcgttgacatgtattgactagttataatgaatca 60
Db 210 gatgtacggccagatatacagcgttgacatgtattgactagttataatgaatca 269
QY 61 attacggggtattatgattacatagcccatatagagttccggttactaactaactagta 120
Db 270 attacggggtattatgattacatagcccatatagagttccggttactaactaactagta 329
QY 121 aatgagccgctgtgctgacccgccaacgaccccccgcattgagctcaataatgacgtat 180
Db 330 aatgagccgctgtgctgacccgccaacgaccccccgcattgagctcaataatgacgtat 389
QY 181 gtcccatagtaacgccaataggaacttccattgacgtcaatagggtagactattacag 240
Db 390 gtcccatagtaacgccaataggaacttccattgacgtcaatagggtagactattacag 449
QY 241 taactgccaacttggcagctacatcaatgtatcatatagccaagtagcccccatttgc 300
Db 450 taactgccaacttggcagctacatcaatgtatcatatagccaagtagcccccatttgc 509
QY 301 gtcaatgacgttaaatggccgctgcatatagccagtaacatgacttggacttt 360
Db 510 gtcaatgacgttaaatggccgctgcatatagccagtaacatgacttggacttt 569
QY 361 cctacttggcagctacatcagttatgltacgtctattaccatggtgactgttttg 420
Db 570 cctacttggcagctacatcagttatgltacgtctattaccatggtgactgttttg 629
QY 421 cagttacatcatggtgctggtatagcgttttactacgaggtattccaagttccacccc 480
Db 630 cagttacatcatggtgctggtatagcgttttactacgaggtattccaagttccacccc 689
QY 481 attgacgtcaatgaggtttgttttggcacaacaatcaacggacttccaanaatgctgt 540
Db 690 attgacgtcaatgaggtttgttttggcacaacaatcaacggacttccaanaatgctgt 749
QY 541 aacaactcgcgcccatatgacgcaaatggcgtagcggtgtacggtggaggtctcata 600
Db 750 aacaactcgcgcccatatgacgcaaatggcgtagcggtgtacggtggaggtctcata 809
QY 601 agcagagctctctgtgctactagagaaacccactgtctactgtgctatcgaaatatacg 660
Db 810 agcagagctctctgtgctactagagaaacccactgtctactgtgctatcgaaatatacg 869
QY 661 actcaactaagagacccaagctgtg----- 687
Db 870 actcaactaagagacccaagctgtgcgcggtgtacacactctcttcgcgcatcgctgtctg 929

QY 688 ----- 687
Db 930 cgaggcgacgtgtgtggctcgcggttgaggagaacactcttcggtcttccagttactc 989
QY 688 ----- 687
Db 990 ttgagtcgaaacccgtcgcttcggaacggttactcgcacagagagactcagagagt 1049
QY 688 ----- 687
Db 1050 ccgcatcgaccggtatcggaacactctcagaggaacttgaataacacagaagttaactgta 1109
QY 688 ----- 723
Db 1110 agttagcttttgccttttatttattcagttccggtacggtgtgtgtaataacaa 1169
QY 724 tgcgtgaattctgagatat----- 743
Db 1170 aactgcttcagtgagagtgcttactctactagcctgtaagagtgtaactctgc 1229
QY 744 ----- 774
Db 1230 tctaagagctcggaatgtgacccgcgctgacgtcagttcagacgaattcgtaagata 1289
QY 775 ctagaagccctattctatagtgtaacctaaatgctgagctcgtatcagctgact 834
Db 1290 tcatgagccctattctatagtgtaacctaaatgctgagctcgtatcagctgact 1349
QY 835 gtgccttcagtttcgagacacactgtgtgttggcccccgcgcttccttcagacctg 894
Db 1350 gtgccttcagtttcgagacacactgtgtgttggcccccgcgcttccttcagacctg 1409
QY 895 gaaggtgccaactccacactgtctcttccataaataatgaggaatgtgacgtcgtctg 954
Db 1410 gaaggtgccaactccacactgtctcttccataaataatgaggaatgtgacgtcgtctg 1469
QY 955 agtaggttcaatctatctctgtgggtgtggttgggcaagacgaaggggagagatttg 1014
Db 1470 agtaggttcaatctatctctgtgggtgtggttgggcaagacgaaggggagagatttg 1529
QY 1015 gaagacaatagcaagcagctggtggga----- 1040
Db 1530 gaagacaatagcaagcagctggtggga----- 1040
QY 1041 ----- 1040
Db 1590 ttccacgcgccttctatgaaggttggtgcttcggaatcgttttcggagcgcgctg 1649
QY 1041 ----- 1040
Db 1650 gatgatcctccagcggtggtatcctcagttgagttcttcgcccacccaactgtttat 1709
QY 1041 ----- 1040
Db 1710 tgcagctataatggttaacaataaagcaatagcatcaacaattccacaataaagcatt 1769
QY 1041 ----- 1040
Db 1770 ttttactgtatctagttgtgtgttgcacaactcaatcatatcatatcatgtctg 1829
QY 1041 ----- 1040
Db 1830 tatacgtcgacctctagtagagcttggtlaatcatgtaacatgcttctgtgtg 1889
QY 1041 ----- 1040
Db 1890 aaattgtatcgcgtcacaattccacaacacatcagcgaggaagcacaatgttaagc 1949
QY 1041 ----- 1062
Db 1950 ctgggtgtcctaatgagtgagctaacatcaattattggtgttggtcctacacgtccgctt 2009

```

OY 1063 -----gagggcgaaagaaaccagctgcatlaatgaatcggccaacgcgcgggagagag 1114
    ||| 1 |||||
Db 2010 ccagtcgggaaacctgctgcgccagctgcatlaatgaatcggccaacgcgcgggagagag 2069
OY 1115 cgglttcgcatctgggagcgtcttcctcgtctcgtctcgtctcgtctcgtctcgtctcgt 1174
    |||||
Db 2070 cgglttcgcatctgggagcgtcttcctcgtctcgtctcgtctcgtctcgtctcgtctcgt 2129
OY 1175 tcggttcgagcgagcgatcatcagctcaactaaaggcggtatagcttatccagaatc 1234
    |||||
Db 2130 tcggttcgagcgagcgatcatcagctcaactaaaggcggtatagcttatccagaatc 2189
OY 1235 aggggataacgcaggaagaaacatgtgagcaaaaggccagcaaaaggccaggaacgttaa 1294
    |||||
Db 2190 aggggataacgcaggaagaaacatgtgagcaaaaggccagcaaaaggccaggaacgttaa 2249
OY 1295 aaaggccggttcgtggttccttcctcaatagctcgcgcgcgcgcgcgcgcgcgcgcgcgc 1354
    |||||
Db 2250 aaaggccggttcgtggttccttcctcaatagctcgcgcgcgcgcgcgcgcgcgcgcgcgc 2309
OY 1355 tcgagcgtcaagtcagaggtgagcgaaacccgcaagagactataagataccagcgcttcc 1414
    |||||
Db 2310 tcgagcgtcaagtcagaggtgagcgaaacccgcaagagactataagataccagcgcttcc 2369
OY 1415 cccctggaagcctccctcgtcgtcctcgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1474
    |||||
Db 2370 cccctggaagcctccctcgtcgtcctcgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2429
OY 1475 cgccttcctccttcgaggaagcgltgagcgttctcgaatgctcaacgctgtagtatccag 1534
    |||||
Db 2430 cgccttcctccttcgaggaagcgltgagcgttctcgaatgctcaacgctgtagtatccag 2489
OY 1535 ttgggtgtagtctgctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1594
    |||||
Db 2490 ttgggtgtagtctgctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 2549
OY 1595 ccgctgagccttaaccgtaactacgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1654
    |||||
Db 2550 ccgctgagccttaaccgtaactacgtctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 2609
OY 1655 gccacgtggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1714
    |||||
Db 2610 gccacgtggcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2669
OY 1715 agcgttccttaagctgtgagccttaactacgctacagcagcagcagcagcagcagcagcagc 1774
    |||||
Db 2670 agcgttccttaagctgtgagccttaactacgctacagcagcagcagcagcagcagcagcagc 2729
OY 1775 cgcctcgtcgaagcagcttaactccttcggaagaaagctgtagctcctgtaaccggaaca 1834
    |||||
Db 2730 cgcctcgtcgaagcagcttaactccttcggaagaaagctgtagctcctgtaaccggaaca 2789
OY 1835 aaccacgcgtcgtgtagcggtgttttttttttttttttttttttttttttttttttttttt 1894
    |||||
Db 2790 aaccacgcgtcgtgtagcggtgttttttttttttttttttttttttttttttttttttttt 2849
OY 1895 agcgttccttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1954
    |||||
Db 2850 agcgttccttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2909
OY 1955 ctcaagcttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2007
    |||||
Db 2910 ctcaagcttaagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 2962

```

RESULT 15

AA23778 standard; DNA; 8705 BP.

XX AA23778;
XX
XX 14-JAN-2000 (first entry)
XX

```

DE Vector pshuttle DNA.
XX Antisense; DNA library; identification: multiple cloning site; MCS;
KW Inhibition; ss.
XX Synthetic.
XX MO9950457-A1.
XX
XX 07-OCT-1999.
XX
XX 28-MAR-1999; 99WO-US06742.
XX
XX 28-MAR-1998; 98US-0079792.
XX 06-NOV-1998; 98US-0107504.
XX
XX (UTAH ) UNIV UTAH RES FOUND.
XX
XX Ruffner DE, Pierce ML, Chen Z;
XX
XX MPI; 1999-610866/52.
XX
XX Production of antisense libraries, used for identifying antisense
XX agents and for identifying target sites for antisense-mediated
XX inhibition of a selected gene
XX
XX Claim 16; Page 43-50; 63pp; English.
XX
XX This invention describes a novel method for generating an antisense
XX library targeted to a selected RNA transcript. The methods can be used
XX for identifying antisense agents and for identifying target sites for
XX antisense-mediated inhibition of a selected gene. The use of a direct
XX library for target site selection significantly simplifies the screening
XX process, since only very small libraries need be prepared and assayed.
XX This sequence represents the vector pshuttle which is used in the method
XX of the invention.
XX
SQ Sequence 8705 BP; 1875 A; 2575 C; 2078 G; 2177 T; 0 other;

Query Match 13.2%; Score 1068; DB 20; Length 8705;
Best Local Similarity 99.9%; Pred. No. 1.4e-150;
Matches 1079; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 gatgtacggcgccagatatacgtctgacatgtatttgactagttatgaatgaatca 60
    |||||
Db 2708 gatgtacggcgccagatatacgtctgacatgtatttgactagttatgaatgaatca 2767
OY 61 attacgggtcatatgatacagccatataatgaggtccggttacataactacgta 120
    |||||
Db 2768 attacgggtcatatgatacagccatataatgaggtccggttacataactacgta 2827
OY 121 aatggccgcgtcgtcgcagccccaacgaccccccgcacattgacgtcaataatgacgtat 180
    |||||
Db 2828 aatggccgcgtcgtcgcagccccaacgaccccccgcacattgacgtcaataatgacgtat 2887
OY 181 gtcccatatgtaacgcaaatgaggttccatctgagcgtcaaatgaggttgaactaacg 240
    |||||
Db 2888 gtcccatatgtaacgcaaatgaggttccatctgagcgtcaaatgaggttgaactaacg 2947
OY 241 taaactgcccacttggcagatatacagagtgatcatatgccaagtacagcccccattagc 300
    |||||
Db 2948 taaactgcccacttggcagatatacagagtgatcatatgccaagtacagcccccattagc 3007
OY 301 gtcaatgagcgttaaatgagccgcgtcgtcgtatagcccaatgacgtacgttggagactt 360
    |||||
Db 3008 gtcaatgagcgttaaatgagccgcgtcgtcgtatagcccaatgacgtacgttggagactt 3067
OY 361 cctaactggcagtaactacagatattagcaatcgcgtacattaccatggtgtagtggtttgg 420
    |||||
Db 3068 cctaactggcagtaactacagatattagcaatcgcgtacattaccatggtgtagtggtttgg 3127
OY 421 cagtatcaatgagcgtgtagtagcggttgtagcagcgggagtttccaagctccacccc 480

```

```
Db 3128 cagtaacatcaatggcgctggaatagcgtcttgaccacagggaattccaaagctccacccc 3187
QY 481 attgacgtcaatggaggtttgttgacacaaatcaacgggactttccaaaatgtcgt 540
Db 3188 attgacgtcaatggaggtttgttgacacaaatcaacgggactttccaaaatgtcgt 3247
QY 541 aacaactccgcccattgaagcaaatggggcgttagcggtgaaggtcctatata 600
Db 3248 aacaactccgcccattgaagcaaatggggcgttagcggttagcggttagcggtcctatata 3307
QY 601 agcagaactctcttggttaacagaagaaccactgctt-actggcttaacgaattaatatc 659
Db 3308 agcagaactctcttggttaacagaagaaccactgctt-actggcttaacgaattaatatc 3367
QY 660 gactcaactataggaggaaccaagcttggtaaccgagctcgatccaactagtaacggcgcc 719
Db 3368 gactcaactataggaggaaccaagcttggtaaccgagctcgatccaactagtaacggcgcc 3427
QY 720 agtgtgtggaattctgcagatalccatcaacactggcgccgctcgagcatgcatctaga 779
Db 3428 agtgtgtggaattctgcagatalccatcaacactggcgccgctcgagcatgcatctaga 3487
QY 780 ggagccatattctatagtgctacccaatgctagaagctcgctgatacagctgactgtgcc 839
Db 3488 ggagccatattctatagtgctacccaatgctagaagctcgctgatacagctgactgtgcc 3547
QY 840 ttctagttgcagacatcgttgtttggccctcccccgtgccttcccttgacccttgaaagg 899
Db 3548 ttctagttgcagacatcgttgtttggccctcccccgtgccttcccttgacccttgaaagg 3607
QY 900 tgccactcccaactgctcttccctaataatgaggaatgtcatcgcatgtctgagtag 959
Db 3608 tgccactcccaactgctcttccctaataatgaggaatgtcatcgcatgtctgagtag 3667
QY 960 gtgtcatctatctctggggggtggggcaggacagcaaggggaggaattgggaaga 1019
Db 3668 gtgtcatctatctctggggggtggggcaggacagcaaggggaggaattgggaaga 3727
QY 1020 caatagcagagcatgtcggggaatgcggtgggtctctatggctctgagcggaagaaccag 1079
Db 3728 caatagcagagcatgtcggggaatgcggtgggtctctatggctctgagcggaagaaccag 3787
```

Search completed: May 24, 2002, 03:46:09
Job time: 13566 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2002, 00:30:03 ; Search time 134.6 Seconds
(without alignments)
14809.182 Million cell updates/sec

Title: US-09-778-516A-1

Perfect score: 8115
Sequence: 1 gatgtacggcgacagatatcac.....gcgtctgcgtacgacggtc 8115

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/prodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/prodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/prodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/prodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/prodata/2/lna/PCUTUS.COMB.seq:*
6: /cgn2_6/prodata/2/lna/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1240.4	15.3	3853	3	US-08-801-092-5
2	1068.6	13.2	4026	3	US-08-801-092-19
3	1053.6	13.0	4326	4	US-08-760-615-7
4	981.8	12.1	6253	3	US-08-893-327-15
5	981.8	12.1	6280	3	US-08-893-327-17
6	981.8	12.1	6280	3	US-08-893-327-19
7	973	12.0	4283	1	US-08-343-401A-3
8	973	12.0	4283	1	US-08-445-265A-1
9	973	12.0	4283	3	US-08-990-442-1
10	968.2	11.9	3987	4	US-09-082-649B-83
11	961.8	11.9	3987	4	US-09-082-649B-84
12	921.8	11.4	6285	1	US-08-467-420A-49
13	921.8	11.4	6285	1	US-08-470-110A-48
14	921.8	11.4	6285	1	US-08-667-769A-49
15	921.8	11.4	6285	2	US-08-940-371-49
16	921.8	11.4	6285	2	US-08-940-371-49
17	921.8	11.4	6285	5	PCT-US95-17082A-49
18	906.4	11.2	3130	4	US-09-038-141-1
19	906.4	11.2	3789	4	US-09-075-019-8
20	906.4	11.2	3918	2	US-08-495-500-2
21	906.4	11.2	4249	3	US-08-801-092-33
22	906.4	11.2	4453	1	US-08-770-761A-4
23	906.4	11.2	4454	2	US-08-738-172-2
24	906.4	11.2	4454	2	US-08-613-861-2
25	906.4	11.2	4540	1	US-08-770-761A-6
26	906.4	11.2	4713	4	US-09-194-285-7
27	906.4	11.2	4724	4	US-09-194-285-8

28	906.4	11.2	4824	2	US-08-485-139-5	Sequence 5, Appl
29	906.4	11.2	4824	3	US-08-750-357-5	Sequence 5, Appl
30	906.4	11.2	4883	1	US-08-064-121-4	Sequence 4, Appl
31	906.4	11.2	4883	1	US-08-318-772A-1	Sequence 1, Appl
32	906.4	11.2	4883	1	US-08-478-015-4	Sequence 4, Appl
33	906.4	11.2	4883	3	US-08-475-975-4	Sequence 4, Appl
34	906.4	11.2	4883	3	US-09-084-889-4	Sequence 4, Appl
35	906.4	11.2	4949	3	US-09-138-024-22	Sequence 22, Appl
36	906.4	11.2	5399	1	US-08-064-121-1	Sequence 1, Appl
37	906.4	11.2	5399	1	US-08-478-015-1	Sequence 1, Appl
38	906.4	11.2	5399	3	US-08-475-975-1	Sequence 1, Appl
39	906.4	11.2	5399	3	US-09-084-889-1	Sequence 1, Appl
40	906.4	11.2	5618	3	US-08-799-569-1	Sequence 9, Appl
41	906.4	11.2	5620	1	US-08-104-072B-9	Sequence 9, Appl
42	906.4	11.2	5620	1	US-08-351-413-3	Sequence 3, Appl
43	906.4	11.2	5620	2	US-09-025-583-3	Sequence 3, Appl
44	906.4	11.2	5642	1	US-08-318-772A-2	Sequence 2, Appl
45	906.4	11.2	5653	1	US-08-073-836-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-801-092-5
; Sequence 5, Application US/08801092
; Patent No. 6074850
GENERAL INFORMATION:
APPLICANT: Antelman, Douglas
APPLICANT: Gregory, Richard J.
TITLE OF INVENTION: Tissue Specific Expression of
TITLE OF INVENTION: Retinoblastoma Protein
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,092
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,517
FILING DATE: 15-NOV-1996
CLASSIFICATION: 514
AUTHOR/AGENT INFORMATION:
NAME: Filts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 016930-001020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 703-576-0300
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3853 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 209..250
FEATURE:
NAME/KEY: CDS

Query Match	15.38;	Score 1240.4;	DB 3;	Length 3853;
Best Local Similarity	74.18;	Pred. No. 1.2e-267;		
Matches 1911; Conservative	0;	Mismatches 96;	Indels 573;	Gaps 4;

[illegible]

Db 1770 GTAACCTGGGGTGCCTATGATGAGCTAACTACATTAAATTTGGTTGGCCTCACTG 1829
Qy 1058 cttctc-----gaagcgaagaagaaccagctgcattatgaatcggcgaagcgg 1107
Db 1830 CCGCTTTTCAGTCGCGGAACTGCTGTCGACCTGATTAAATGAAATGCGCAACGGCGG 1889
Qy 1108 ggaagaagcggcttgctgctatggcgccttcgccttcgcctgcctgcctgcctgc 1167
Db 1890 GGAAGAGCGGTTTGGCTATTGGCGCTCTTCCGCTTCTCGCTCACTGACTGCTCGCTG 1949
Qy 1168 cggctcgttcgcctgc 1227
Db 1950 CGGCTGCTGGCTGGCGGAGCGGTATCAGCTCACTCAAAAGGGGATTAATAGGTTATCCA 2009
Qy 1228 cagaatcagaagataacgaagaagaacaatgtgaagcaaaagccagcaaaagccagga 1287
Db 2010 CAGAAATCAGGGGATTAACGAGGAAACATGTGACCAAAAGGCCAGCAAAAGGCCAGGA 2069
Qy 1288 accgtaaaagcgcgcgttcgcgcgttccttcataagctcgcgcgcgcgcgcgcgcgcgc 1347
Db 2070 ACCGTAATAAGCGCGGTTGCTGCGCTTTTTCATAGGCTCCGCCCTCGACGAGCATC 2129
Qy 1348 acaaaaaatcgaagcgtcaagtcagaagtggaagaccgaagagactataagaataccag 1407
Db 2130 ACAAAAATCGACGCTCAAGTCAAGAGTGCGGAAACCCGACAGAGCATATAAAGATACAGG 2189
Qy 1408 cgttcccccctggaagctccctcgctgcctcctcgttcgcgcgcgcgcgcgcgcgcgc 1467
Db 2190 CGTTTCCCTCCCTGGAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2249
Qy 1468 accgtcgcgccttcctccctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1527
Db 2250 ACTGTGCGCCCTTTTCTCCCTTGGGGAAGCGTGCGCTTCTCATATCTCAAGCTGTAGGT 2309
Qy 1528 atctcaagtcggtgtagtgcgttcgcctcaagctggcgtggtgcgcgcgcgcgcgcgc 1587
Db 2310 ATCTCAGTTGCGGTGTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2369
Qy 1588 agcccgagc 1647
Db 2370 AGCCCGACCGCTGCGCTTATCCGTAATGCTGTGAGTCCAAACCGGTAAAGCACAG 2429
Qy 1648 acttatcgccactgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 1707
Db 2430 ACTTATCGCCACTGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2489
Qy 1708 gtgctacagagtccttgaagtggtggtcctaactacgcgtacactagaagagacagta 1767
Db 2490 GTGCTACAGAGTCTTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2549
Qy 1768 gtaactgcgcctgctgtaagcagtaacttcggaaaaagagtgtagctcttgatcgg 1827
Db 2550 GTATCTGCGCTGTGCTGAAGCAGTACCTTCGAAAAAGAGTGTGCTGCTGCTGCTGCTG 2609
Qy 1828 gcaaaacaacacacgcgtgtagcgtggtgtttttgtttgcaagcagcagatgaagcga 1887
Db 2610 GCAAAACAACACACGCTGTAGCGGTGTTTTTTTGTTCAGACGACGAGATTACGCGCA 2669
Qy 1888 gaaaaaaagatctcaagaagatccttgaatcttctacgcgggtcgaagcagtgga 1947
Db 2670 GAAAAAAAGATCTCAAGAAATCTTTGATCTTTTCTACGGGGCTGACGCGTCACTGGA 2729
Qy 1948 acgaaaaacacagtgtaagagatttggtcacaagcagatacatattgaatgattaga 2007
Db 2730 ACGAAAACTCACGTTAAGGATTTTGGTATGATGATTATCAAAAGATCTTCACCTAGA 2789

RESULT 2

US-08-801-092-19
Sequence 19, Application US/08801092

Patent No. 6074850
GENERAL INFORMATION:

APPLICANT: Antelman, Douglas
APPLICANT: Gregory, Richard J.
TITLE OF INVENTION: Tissue Specific Expression of
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,092
FILING DATE: 14-FEB-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/751,517
FILING DATE: 15-NOV-1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Fitts, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 016930-001020
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 703-576-0300
INFORMATION FOR SEQ. ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 4026 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 209..250
FEATURE:
NAME/KEY: CDS
LOCATION: 254..289
FEATURE:
NAME/KEY: CDS
LOCATION: 293..505
FEATURE:
NAME/KEY: CDS
LOCATION: 509..514
FEATURE:
NAME/KEY: CDS
LOCATION: 518..520
FEATURE:
NAME/KEY: CDS
LOCATION: 524..658
FEATURE:
NAME/KEY: CDS
LOCATION: 662..691
FEATURE:
NAME/KEY: CDS
LOCATION: 695..748
FEATURE:
NAME/KEY: CDS
LOCATION: 752..781
FEATURE:
NAME/KEY: CDS
LOCATION: 785..829
FEATURE:
NAME/KEY: CDS
LOCATION: 833..862

Db	3584	GGTAAATACGGTTTCCACAGAAATCAGGGGTAACCGCAGGAAGAAACATGTGACAAAG	3643
QY	1271	ccagcaaaagcccaaggaaccggtataaaagccgcgtgtgtcgtgtcgttltccataagctcgc	1330
Db	3644	CCACCAAAAGCCAGGAAACCGTAAAAAGCCCGCGTTGCGCGGTTTTCATNAGGCTCCG	3703
QY	1331	ccccctgagagcatcacaaaatacgagctcaagtcaagttaggtagggtaaacccgaacag	1390
Db	3704	CCCCCTGACGAGCATCACAAAATGAGCGCTCMACTCAGAGGTGGCAACCCGACAGG	3763
QY	1391	actataaaga taccaggcgttllccccctggaaagctccccctcgcgctcctcctgttcgcac	1450
Db	3764	ACATAAAGATACAGGCGTTTCCCTCGGAAGCTCCCGTGCGCTCCGTTCCGAC	3823
QY	1451	ccctgcgcttaacggatacctgtcgcgttctcccttcgggaagcgtggcgcttctca	1510
Db	3824	CTTGCCCTTACCGGATCCTGTGCCCTTTCTCCCTTGCGGAACGTTGGCGCTTTCGA	3883
QY	1511	atgtcaacgtttagatcatcttagtctcgtttagtgcgttcgtctcaagctggggtgtgt	1570
Db	3884	ATGCTCAAGCTGTAAGGATCTCAATTCGGGTGTAAGGCTGTTCTCTCAAGCTGGGCTGTGT	3943
QY	1571	gcaagcaaccccccglttaagccacacgcgtgcgcgcctatccggttaactcgtctgaatc	1630
Db	3944	GCAGAAACCCCCGTTTACGCGCAGCGCGTGCGCTTATCCGTAATATGCTTGAATC	4003
QY	1631	caaccgcgtatagaacgaactatcgccactggcagcagccacttgg laaacagattatgacg	1690
Db	4004	CAACCCCGTAAAGACGACTTATTCGCGACCTGGCAGGACCACTGGTAACAGATTATGACG	4063
QY	1691	agcaggtatgtagtagcggtgcctcagaagttcttgaagtgtagcctlaactaagctaac	1750
Db	4064	AGCGAGTATGTAGCGGTCCTACAAAGTTCTTGAAGTGTGTGGCTTAACCTACGCGCTAAC	4123
QY	1751	tagaaggaacagiatlttgtglatctgcgtctgtctgtgaagccaagtctaccttcggaaaagat	1810
Db	4124	TAGAAGACAGTATTGTGATCTGCGCTCGTGGTGAAGCAGATTACTTCGGAAGAAAGGT	4183
QY	1811	tgtgtagctctttagtcgcggcaacaacaacacacgcgctgtagcgttgttlttltgtttgcaa	1870
Db	4184	TGTGATGCTTGTATCCGGCAAAACCAACCCCTGTGTACCGGTGTTTTTGTGTGCA	4243
QY	1871	gcaacagattacgcgcagaaaaaagagactcaagaagaatcccttgatccttlttcacggg	1930
Db	4244	GCAACATATTACCGCGCAAGAAAAAAGAGTCTAAGAAGATCTTTGATCTTTTCTACGGG	4303
QY	1931	gtctgacgtctcagltgaaacgaanaactcacgtttaaggaatttgtgtcatgacgcgatacat	1990
Db	4304	GTCGACGCTCATGTGAAACAAACACTCAGTTAAGGATTTTGGTCATGAGATTATACAA	4363
QY	1991	atttgaaatgtatttaga 2007	
Db	4364	AAGGATCTTACCTAGA 4380	

RESULT 6
US-08-893-327-19
Sequence 19, Application US/08893327
Patent No. 6020192
GENERAL INFORMATION:
APPLICANT: Zolotukhin, Sergei
APPLICANT: Hauswirth, William W.
APPLICANT: Muzyczka, Nicholas
TITLE OF INVENTION: Humanized Green Fluorescent Protein
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: USA
ZIP: 77210-4433

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/893.327
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/588.201
FILING DATE: 18-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Kitchell, Barbara S.
REGISTRATION NUMBER: 33,928
REFERENCE/DOCKET NUMBER: DFLA:062\KIT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (713) 789-2679
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 6280 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: 988..1728
US-08-893-327-19

```

Query Match	12.1%: Score 981.8; DB 3; Length 6280;
Best Local Similarity	86.9%; Pred. No. 7,66210;
Matches 1162; Conservative	0; Mismatches 37; Indels 138; Gaps 2
QY 809	ctagaagctcgtatcaagccttcgaactgtgccttcaagtatgcaagcattcgtttggc 868
Db 3044	CTAGAGCTCGGTGATGACCCCTGACCTGATGGCTTTCAGTTCACGACCATGTGTGGTTC 3103
QY 869	cctcccccgttccttccttccttcgaacctcgggaagtgccacacccacitcccttcctaataa 928
Db 3104	CTCTCCCGGTGCTTCTTCCCTTCGTGACCTCGGAAGGTGCCACTCCACTGCTCTTCTTAATTA 3163
QY 929	atgaggaattgcatcgcattgtcgtgaatgaatgtgcatctatcttcgaggagtgaggatgg 988
Db 3164	ATGAGGAATTCATCCATCCGATGTCTGAACTAGTATGATTCATTCCTTGAGGGGCTGGGTGG 3223
QY 989	ggcagagaaacgaagaagggaagatctggagaacaaataagaagagatcttggggg----- 1040
Db 3224	GGCAGGACGACGAGGGGAGGAGTGTGGAGACAAATAGCAGGCATGCTGGGGAGACATCTTA 3283
QY 1041	----- 1040
Db 3284	GGAACCCCTAGTAGATGAGATTGGCCACTCCCTCTCGCGCGCTCGCTCGCTCACTGAGGC 3343
QY 1041	-----tgcgatgggctctatbtgcttcttgaaggcggaagaacag-- 1079
Db 3344	CGCCCGGGCAAAAGCCCGGGCGTGGGCGACCTTTGTGTGGCCCGGCGCTCACTAGACGAGGG 3403
QY 1080	-----ctgacattaatg 1090
Db 3404	AGCGGCGAGAGGGAGTGGGCCAACCCCCCCCCCCCCCCCCCTCGAGCCCTTCATTAATG 3463
QY 1091	aatcgccaacgcgcgagagagagcggtttgcgtattggcgctcttcgcgtcttcgcgt 1150
Db 3464	AATCGGCCCAACGCGGGGGAGAGGGGCGTGTGGTATTTGGGCGCTTCTTCGATCTCTCGGT 3523
QY 1151	cactgaactcgtctgcgtcgttcgttcgtcgtcgtgagagagatcaagctcaactcaaaagc 1210
Db 3524	CACGTACTCTGCTGCGCTTGGGTGTTCCGGCTGGCGGAGCGGATATCAAGCTCACTCAAAGC 3583
QY 1211	ggtataaggttatccacagaatcagggagataacgcaggaanaaacatgttgcaaaaag 1270

Db 3584 GGTAATACGTTATCCACAGATCAGGGGATACGACGAAAGAAACATGTGACCAAAAGG 3643
Qy 1271 ccagcaaaagccaggaacggtlaaaagccggtctgctgctgttttccatagctccg 1330
Db 3644 CCACCAAAAGCCAGGAACCTAAAGGCGCGCTTCCTGCGCTTTTTCATAGGCTCCG 3703
Qy 1331 cccccgcagcagatcacaaaatcgacgctcaatcaaggtgagcgaaccccgacag 1390
Db 3704 CCCCCGTGACAGCATCACAAAATCGACGCTCAAGTCAGAGTGCGCAACCCGACAGG 3763
Qy 1391 actataaagatcacaggcggttcccccgtgaagctccctcgctgctctcgttccgac 1450
Db 3764 ACTATAAAGATACAGGCGGTTCCTCCGTAAGCTCCCTGCGCTCTGCTGTTCCGAC 3823
Qy 1451 ccttcgcttaacggatcacctgctgcgcttctcccttcgggaagcggtgagctttcca 1510
Db 3824 CTTCCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGGAAGCGGCGCTTTCGA 3883
Qy 1511 atgctcagctgtagatctcagttcggtgtagtctgctcccaagctgggctgtgt 1570
Db 3884 ATGCTCAGCGTGTAGTATCTCAATTCGTTAGTGTAGTCTTCGCTCCAAAGCTGGGCTGTGT 3943
Qy 1571 gcacgaaccccccgcttaacgcccgcgctgcgcttaccgtaactatcgltcttgatc 1630
Db 3944 GCACGAACCCCCCTTACGCGCCGACCGCTGCGCTTATCCGTAACATCTGCTTGAATC 4003
Qy 1631 caaccgcgttaagacagactatcgccactgacgacgacgacgctgttaacgaattagag 1690
Db 4004 CAACCCGTTAAGACACCTTATGCGCACTGCGACAGCCACTGTAACAGGATTTAGCAG 4063
Qy 1691 agcgaaggtatgtagcggtgtacagagttctgaagtggtgacctaaactacgacgac 1750
Db 4064 AGCGAGTATGTAGCGCGGTCTACAGAGTCTTGAAGTGTGGGCTTACTACGCGGTACAC 4123
Qy 1751 tagaagaacggtatttggtatctgctgctctgctgaagcgaattacctcggaaaaaggt 1810
Db 4124 TAGAAGGACAGTATTTGTTATCTGCGCTGCTGCTGAAGCAATTAACCTTCGGAAGAGGT 4183
Qy 1811 tggtagctctgtgacgcgcaacaacaacacgctgtagcggtgtttttgtttgtaa 1870
Db 4184 TGGTAGCTCTTGATCCGCAAAACACCGCTGTTAGCGGTGTTTGTGTTGAA 4243
Qy 1871 gcacgagattacgcgcgaaaaaaaggatctcaagaagatcccttgatcttctacgag 1930
Db 4244 GCACGACATTAACGCGCAAGAAAAAGATCTCAAGAGATCTTGTGATCTTTCACGGG 4303
Qy 1931 gtcggaagctcagtggaagcaaaactcaactcaaggtatggtatgtagcgagatacat 1990
Db 4304 GTCTGAGGCTCAGTGGAACGAACACTACGTTAAGGGATTTGGTCATGAGATTAACAA 4363
Qy 1991 attggaatgtaattaga 2007
Db 4364 AAGGATCTTCACCTAGA 4380

RESULT 7

US-08-343-401A-3
Sequence 3, Application US/08343401A
Patent No. 5661132

GENERAL INFORMATION:

APPLICANT: Swain, William F
APPLICANT: Macklin, Michael D
APPLICANT: Eriksson, Elov
APPLICANT: Andree, Christophe
TITLE OF INVENTION: Improved Wound Healing
NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: Quarles & Brady
STREET: PO Box 2113
CITY: Madison
STATE: WI
COUNTRY: USA
ZIP: 53701-2113

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/343,401A
FILING DATE: 22-NOV-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J
REGISTRATION NUMBER: 27,386
REFERENCE/DOCKET NUMBER: 11-229-9103-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 608-251-5000
TELEFAX: 608-251-9166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 4283 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: pMRG1630
FEATURE:
NAME/KEY: exon
LOCATION: 713..721
FEATURE:
NAME/KEY: exon
LOCATION: 981..1253
FEATURE:
NAME/KEY: CDS
LOCATION: join(713..721, 981..1253)
FEATURE:
NAME/KEY: sig-peptide
LOCATION: 713..1049
US-08-343-401A-3
Query Match 12.08; Score 973; DB 1; Length 4283;
Best Local Similarity 83.28; Pred. No. 6.1e-208;
Matches 1216; Conservative 0; Mismatches 30; Indels 215; Gaps 1;
Qy 762 ctgagacatgcatcagaagggcccatatctatagtgtaacctaaatgctagagctgctg 821
Db 1277 CTCGAGCATGCATCTAGAGGGCCCTATTTCTATAGTGTACCTAAATGCTAGAGCTCGTG 1336
Qy 822 atcagcctcgaactgacctctagttgccaagcatalgtgtttgccccctcccgctgc 881
Db 1337 ATCAGCCCTCGACTGTGCTTCTAGTGGCAGCATCTGTGTTGCCCTCCCGCTGCC 1396
Qy 882 ttccctgaacctggaaggtgccaactccactgctcttccctaataaatgaggaattgc 941
Db 1397 TTCCCTTGACCTCGAAGGTGCGACCTCCACCTGCTCTTCTTAATTAATGAGGAATTGC 1456
Qy 942 atcgcatgtctgagtagtgltcatctatctctgaggggtgaggtgagggcgagagcaa 1001
Db 1457 ATCGCATTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1516
Qy 1002 ggggagagattgggaagacaatagcagagcatgctgaggtgaggtgaggtgaggtgag 1061
Db 1517 GGGGAGAGATTGGGAAGACATAGCAGGAGCATGTGGGAGTGGGCGGTCTGATGAAAC 1576
Qy 1062 tgaggcggaaga----- 1074
Db 1577 AGCTGGGCTGAGCATGCAAGCTTGACTATTCTATAGTGTACCTAAATAGCTTGGCGT 1636
Qy 1075 ----- 1074
Db 1637 AATCATGCTCATAGCTGTTTCTGTGTGAATTTGTTATCCGCTACAAATTCACACACAA 1696
Qy 1075 ----- 1074

```

Db 1697 TACGAGCCGGAAGCATAAAGTAAAGCCTGGGTCCTAATGATGAGCTAATCACAAT 1756
Qy 1075 -----acagctgcat 1086
Db 1757 TAATTGGCTTGGGCTCACTGCCCGCTTTCCACTCGGGAACCTGTCGCTCAGCTGCATT 1816
Qy 1087 aatgaatcgccaaagcgcgagagagagcggttgcgtatctggcgctctcccgcttcc 1146
Db 1817 AATGAATCGGCCAACGCGGGGAGAGCGGTTGGGTATGCGGCTCTCCGCTTCCCT 1876
Qy 1147 cgctactgactcgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1206
Db 1877 CGCTACTGACTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1936
Qy 1207 aagcggtataagcgttatccacagaaatcagggagataacgcagaaagaaacgttgagaa 1266
Db 1937 AGCGGCTATATAGCGTTATCCACAGAAATCAGGGGATTAACGAGAAAGAAATGTGAGCA 1996
Qy 1267 aagcgcaagaaaggcgagaaacgtaaaaggcgcggttgcgtgctgctgcttccataggc 1326
Db 1997 AAGGCGAGCAAAAGGCCAGGAACCGTAAAGGCGCGCTTGCCTGGCGTTTTCGATAGGC 2056
Qy 1327 tccggcccccctgacagagacacaaatcgagcgtcgaagtcaaggtgagcgaaaccga 1386
Db 2057 TCCGCCCCCTGACGAGCATCAAAAATCGAGCTCAAGTCAAGAGTGCGCAAAACCCGA 2116
Qy 1387 caggaactaaagataaccagcgcttccccctgagaaagctccctgctgctcctcgttcc 1446
Db 2117 CAGGACTATTAAGATACGAGCGGTTCCCGCTGGAAGCTCCGCTGCGCTCTCCGTTTC 2176
Qy 1447 cgaacctgcgcttccaggaatacctgctgctgcttccctcccggaagcgctgctt 1506
Db 2177 CGACCCCTGCGCTTACCGGATACCTGCTCGCCCTTCTCCCTCGGGAACGAGCGCGCTTT 2236
Qy 1507 ctcaatgctcaagcgtgtaagtaaccagtgctgctgctgctgctgctgctgctgctgct 1566
Db 2237 CTCATAGCTCAGCGCTGTAAGTATCTCAGTTCGCTGAGTCCGCTTCCCTCAAGCTGAGCT 2296
Qy 1567 gctgcaagcaaccccccgcttccagccagcgtgctgcttcccttcccgtaactaactgct 1626
Db 2297 GTGTGACGAAACCCCGCTTACGCCGACCGCGCTGCGCTTATCCGATACGATCAGCTTG 2356
Qy 1627 agtccaacccgtaagaaacagctatcgccactgagcagcaacacactgtaacagatta 1686
Db 2357 AGTCCAAACCCGCTAAGACACGACTTATCCCACTGGCAGCAGCCACTGGTAACAGGATTA 2416
Qy 1687 gaagagcgagatgtagcggtgctacagagttcctgaagtgtgctgctgctgctgctgct 1746
Db 2417 GCAGAGCGAGGTATGTAGCGGTGCTGACAGAGTTCTGTGAAGTGTGCGCTTAACGAGCT 2476
Qy 1747 aactagaagaagatattgtatctgctgctgctgctgctgctgctgctgctgctgctgct 1806
Db 2477 ACACTAGAAGGACAGATTTGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2536
Qy 1807 gacttggtagctctgattcccgcaaaacacacacacacacacacacacacacacacacac 1866
Db 2537 GAGTTGGTAGCTCTTGTATCCGCAACACACACACACACACACACACACACACACACACAC 2596
Qy 1867 gcaacagcaagatacgcgcagaaagaaagatactcaagaagatcccttgatcttctta 1926
Db 2597 GCAACGACAGATTTAGCGGCAAAAGAGATCTCAAGAAATCTTTGATCTTTTCTTA 2656
Qy 1927 cggggtctgacgctcaagtgaaacgaacacacacacacacacacacacacacacacac 1986
Db 2657 CGGGGTCTGACGCTCAGTGAAGAAACTCAGCTTAAGGAGTTTGGTCAATGAGATTAAT 2716
Qy 1987 acatattgaaatgatttga 2007
Db 2717 CAAAAGAGATTTCACTAGA 2737

```

```

US-08-445-265A-1
; Sequence 1, Application US/08445265A
; Patent No. 5697901
; GENERAL INFORMATION:
; APPLICANT: Eriksson, Rlof
; TITLE OF INVENTION: GENE DELIVERY BY MICRONEEDLE INJECTION
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Quarles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/445,265A
; FILING DATE:
; CLASSIFICATION: 604
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 110229, 91080
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 608-251-5000
; TELEFAX: 608-251-9166
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "Plasmid DNA"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(713..721, 981..1250)
; US-08-445-265A-1

Query Match 12.0%; Score 973; DB 1; Length 4283;
Best Local Similarity 83.2%; Pred. No. 6,1e-208;
Matches 1216; Conservative 0; Mismatches 30; Indels 215; Gaps 1;

Qy 762 ctcagacatgcatctagaagggccctattctatagtgtaaccctaaatgtagagctgctg 821
Db 1277 CTCGACATGCATCTTAGAGGGCCCTATTCTATAGTGTCACTAAATGCTAGAGCTCGCTG 1336
Qy 822 atcagcctcgactgctgcttctagtgccagccatcgttgttgcctccctcccgctgcc 881
Db 1337 ATCAGCCCTGAGACTGTGCCCTTCTAGTTGCCAGCAGCATCTGTGTTGGCCCTCCCGCTGCC 1396
Qy 882 ttccttgacctggaagtgccacatcccaactgcttcccttaataaagaggaatttc 941
Db 1397 TTCCTTGACCCGGAAGGTGCCACCTCCACTGCTCTTCTTAATAAATGAGGAAATTGC 1456
Qy 942 atcgacatctcgagtagtgatctatctcttgagggttgagggtgagagcagagca 1001
Db 1457 ATCGCATCTGTGAGTAGGTGTCAATCTATTCTGGGGGTGGGTGGGCGAGGACAGCA 1516
Qy 1002 ggggagagatlgggaacaatagcagagatcgtgagatgctgagctctatgcttc 1061
Db 1517 GGGGAGAGTATGGGAAGAATAAGCAGGATCTGGGATGGGTGGCTGTATGGAAC 1576
Qy 1062 tgaagcggaaga----- 1074
Db 1577 ACCTGGGCTCGAGCATGCAAGCTTGAGTATCTATAGTGCACCTAATAATAGCTGGCGCT 1636
Qy 1075 ----- 1074

```


|||||
Db 4549 CGCCCCCTGACGAGATCAAAAAATCGACGCTCAAGTCAAGAGTGGCCAAACCCACA 4608
Qy 1389 ggcataaagataccagcggttccccctggaaagctccctcgtcgtcctcgttcog 1448
Db 4609 GGAATATAAAGATACCAAGCGGTTTCCCGCTGGAAGCTCCCGTGGCTCTCTGTTCCG 4668
Qy 1449 accctgcgtttaccgataacctcgtctccgtcttcccttcctcggaaagcggtgcttctc 1508
Db 4669 ACCCTGCCGCTTACCGGATACCTGTCCGCTTCTCCCTTCCGGAAACGTTGGCCCTTCT 4728
Qy 1509 caatgcacagctcgttaagatctcagctcgttgaagtcgttcgtcccaagctggtctgt 1568
Db 4729 CAATGCTCACGCTGTAGTATCTAGTTCGGTGTAGTGTGCTGCTCCCAAGCTGGCTGT 4788
Qy 1569 gtgcagacacccccgttcagcccgacgctcgtcgttcccttcogtlaactcgtctgag 1628
Db 4789 GTGCACGACACCCCCGTTTCAAGCCCGACCGCTGCTTATCCGGTACTATCTGCTTGA 4848
Qy 1629 tccacccggttaagacagactatgcgaactggcgaagcgaactgtgtaacagattagc 1688
Db 4849 TCCACCCGGTAAACACGACTTATCGCACTGGCAGACGCACTGTATACAGGATTAGC 4908
Qy 1689 agagcagatgtagcagctcgtcagagatctcgtgaagtcgtcgtcctaaactagcgttac 1748
Db 4909 AGAGCGAGTATGTAGGCGGTGCTACAGAGTCTTGAAGTGTGCTTAACTACGCTTAC 4968
Qy 1749 actagaagacagattgtglatctcgtctcgtcgtgaagcagttacccctcggaaaga 1808
Db 4969 ACTAGAAGACAGTATTTGTATCTGCGCTCTGCTGAAGCAGATTACCTTCGGAAGAA 5028
Qy 1809 gtgtgagctcttgatccggcaaacacacacgctgtagcggtgtgtgtgtgtgtgtgtgc 1868
Db 5029 GTTGATGCTCTTGTATCCGCAACAAACACCGCTGTGAGCGGTGTGTGTGTGTGTGTGC 5088
Qy 1869 aagcagacatcagcgcagaaagaaagatctcaagaagatccttgatcttctcagc 1928
Db 5089 AAGCAGAGATTACGCGCAGAAAAAAGATCTCAAGAGATCTTGTGATCTTTCTTACG 5148
Qy 1929 ggtgtgagctcagtgtagcagaaactcaactcaaggaattgtgtcagagcaggttac 1988
Db 5149 GGGTCTACGCTCAGTGAAGCAAACTCAAGTTAAGGATTTGTGATGATGATTTATCA 5208
Qy 1989 atattgatactattaga 2007
Db 5209 AAAAGATCTTCACTAGA 5227

RESULT 13
US-08-470-110A-49
; Sequence 49, Application US/08470110A
; Patent No. 5693323
; GENERAL INFORMATION:
; APPLICANT: Ames, Robert S.
; APPLICANT: Appelbaum, Edward R.
; APPLICANT: Chaiken, Irwin M.
; APPLICANT: Cook, Richard M.
; APPLICANT: Gross, Mitchell S.
; APPLICANT: Holmes, Stephen D.
; APPLICANT: McMillan, Lynette J.
; APPLICANT: Theisen, Timothy W.
; TITLE OF INVENTION: Recombinant IL5 Antagonists Useful in
; TREATMENT OF ALLERGIC DISORDERS
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corp./Corporate
; ADDRESSEE: Intellectual Property
; STREET: P. O. Box 1539-UM220
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,110A
; FILING DATE:
; CLASSIFICATION: 426
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/363131
; FILING DATE: 23-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Sutton, Jeffrey A.
; REGISTRATION NUMBER: 34,028
; REFERENCE/DOCKET NUMBER: P50282
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610 270-5024
; TELEFAX: 610 270-5090
; INFORMATION FOR SEQ ID NO: 49:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6285 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
US-08-470-110A-49

Query Match 11.4%; Score 921.8; DB 1; Length 6285;
Best Local Similarity 82.9%; Pred. No. 1.9e-196;
Matches 1160; Conservative 0; Mismatches 27; Indels 212; Gaps 1;

Qy 821 gatcagctgactgtgctctcctcagttgccaagcactgtgtgtgtgccccctcccggtgc 880
Db 3829 GATCAGCCTTGACTGTCTCTTGTAGTGGCCAGCATCTGTGTGTGCTTCCCTCCGCTGC 3888
Qy 881 ctccctgacccctggaaagtgccactccactgtccttcccttaataatgagaaattg 940
Db 3889 CTTCCTTGACCTCGAAGAGTCCACTCCACTGTCTTCTTAATAAAGAGAAATTG 3948
Qy 941 catcagctgtcgtgagtagtgcattctatctctggtgggtggtggtggtggtggtggtg 1000
Db 3949 CATCGCATCTCTGAGTAGTGTGATCTATCTCTGCGGGGTGGGCTGGCAGACAGCA 4008
Qy 1001 agggggagagattggaaagaaatagcagagatcgtggagtggtgggtcctcattgctt 1060
Db 4009 AGGGGAGAGATTGGGAAGCAATAGCAGCATGCTGGGATGGGCTCTATGGAAC 4068
Qy 1061 ctgagcggaaga----- 1074
Db 4069 CAGCTGGGGCTCGATCGAGTATGATCGGGCCGCGATCCGTCGAGACTTGGCGTAA 4128
Qy 1075 ----- 1074
Db 4129 TCATGTGATAGCTGTTTCTGTGTGAATTTATTCCTCCCTCAAAATTCACACAACTA 4188
Qy 1075 ----- 1074
Db 4189 CGAGCCGGAAGCATTAAGTAAAGCTGGGGTGCCTTAATGAGTAGCTACTACATTA 4248
Qy 1075 ----- 1074
Db 4249 ATTGCGTTGGCTGATCGCCCGCTTCCATCGCGGAACCTGTCTGCGCAGCTGCAATTAA 4308
Qy 1089 tgaatggccaacgcgcggggaagagcggttgcgtatggtggtccttcctcctcgtc 1148
Db 4309 TGAATGCGCCAAACGCGGGGAGAGGGGCTTTCGTAATGGGCGCTTCCGCTTCTCG 4368
Qy 1149 ctcaactgactcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtc 1208
Db 4369 CTCACTGACTGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4428
Qy 1209 gcggtatacgtttatccacagatacaggggataacgcaggaagaacatgtgagcaaaa 1268

```

Db 4429 GCGGTAATACGGTTATCCACGAATCAGGGATTAACGAGGAAGAACATGTGAGCAAAA 4488
Qy 1269 ggcagcaaaaagcgagcaaacgtaaaaaagccgctgttcgagcgttttccataagcgc 1328
Db 4489 GGCACGCAAAAAGCGCAGAAACCGTAATAAAGCCCGCTTGCGGCTTTTTCATAGGCTC 4548
Qy 1329 cgccccctgacgagcatcacaaaaatcgacgtcctaagtcagagtgagcgaaaccgaca 1388
Db 4549 CGCCCCCTGACGAGCATCAAAAAATCGACGTCAAGTCAAGAGTGCGCAAAACCCGACA 4608
Qy 1389 ggaactataaataccaggggtttcccccctggaagctcccccgtgcgtccctgttccg 1448
Db 4609 GGACTATAAAGATACCGAGCGCTTCCCTGGAGCTCCCTGAGCGCTCCCTGCTTCCG 4668
Qy 1449 accctgcgcgttacgagatactgtccgccttctccctcgggaagcgtgagccttct 1508
Db 4669 ACCCTGCGCGTTACCGGATACCTGTCCGCCCTTCTCCCTTGGGAGAGTGCGGCTTTCT 4728
Qy 1509 caatgtcgaagctgtagtalcagtcagtcagtcagtcagtcagtcagtcagtcagtc 1568
Db 4729 CAATGCTCACGCTGTAGGTATCTAGTTCGGGTAGGCTGCTCCCTCAAGCTGGGCTGT 4788
Qy 1569 gtgacgaacccccgttcagccgacgcgctgcgcttaccgtaactacgtcttgag 1628
Db 4789 GTGACGAAACCCCGCTTCACGCCGACCGCTGCGCTTATCCGGTAACTATGCTTTCAG 4848
Qy 1629 tccaaaccggttaagacacgacttacgcacactgagcagcagcagcagcagcagcagc 1688
Db 4849 TCCAAACCCGTTAAACACACGACTTATCCGCACTGGCAGCAGCAGCAGCAGCAGCAGC 4908
Qy 1689 agagcagagtagtagcagcgtgtctacagagtccttgaagtgagtgagtgagtgagtc 1748
Db 4909 AGACCGAGGTATGTAGCGGCTGCTACAGAGTTCCTGAAGTGCTGCTCAACTAGCGCTAC 4968
Qy 1749 actagaaggaagatattgtgtagtcgtctgtctgtagcgaacttactctggaagaaaga 1808
Db 4969 ACTAGAAGGAGATATTGATCTGCGCTGTCTGTAAGCCAGTATACCTTGGAAGAAAGA 5028
Qy 1809 gttagtagctctgtatccgcaaaacacacccgctgtgtagcgtgtgtgtgtgtgtgt 1868
Db 5029 GTTGTGAGCTCTGTATCGCGCAACCAACACCGCTGTGAGGTTTGTGTTTGTGTTGC 5088
Qy 1869 aagcagcagatatacgcgcagaaaaaagagatcacaagaatccttgcatttctacg 1928
Db 5089 AAGCAGCAGATTAACGCCACGAAAAAAGAGATCTCAGAAAGATCCTTGATCTTTCTACG 5148
Qy 1929 gggctcagcgtcagtgagcaaaactcaacgtttaagagatttgatcagagcagagtac 1988
Db 5149 GGGCTCAGCCTCAGTGGAGCAAAACTCAGCTTAAGGATTGTCATAGATTATCA 5208
Qy 1989 atattgaatgattaga 2007
Db 5209 AAAAGATCTTCACCTAGA 5227

```

```

ADDRESS: SmithKline Beecham Corp./Corporate
STREET: P.O. Box 1539-0W2220
CITY: King of Prussia
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/667,769A
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/17082
FILING DATE: 22-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/470110
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/467420
FILING DATE: 06-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/363131
FILING DATE: 23-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Sulton, Jeffrey A.
REGISTRATION NUMBER: 34,028
REFERENCE/DOCKET NUMBER: P50503
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5024
TELEFAX: 610-270-5090
INFORMATION FOR SEO ID NO: 49:
SEQUENCE CHARACTERISTICS:
LENGTH: 6285 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
US-08-667-769A-49
Query Match 11.4%; Score 921.8; DB 1; Length 6285;
Best Local Similarity 82.9%; Pred. No. 1,9e-196;
Matches 1160; Conservative 0; Mismatches 27; Indels 212; Gaps 1;
Qy 821 gatagactgactgttccttctagtgccagcagcatctgtgtttgcccctcccccgtgc 880
Db 3829 GATCAGCCTCGACTGCTCCCTTCTAGTGGCCAGCCATCTGTTGGCCCTCCCTCCCGCTGC 3888
Qy 881 ctccctgacccttgaaagtgtccactcccaactgtccttctcctaataaatgagaaatg 940
Db 3889 CTTCCTTGACCTGTGAAGTGCCACTCCACTGTGCTTCTCTATAAAGAGGAATTG 3948
Qy 941 catcgcatgtctgagtagtgatcattcttctgggggttggggttgaggcagagca 1000
Db 3949 CATCGCATGTGCTGAGTAGTGTCATCTATTCGGGGGTGGGGTGGGCGACAGACA 4008
Qy 1001 agggggagagattgggaagaataagcagcagatctggagtgagtgagtgagtgagtc 1060
Db 4009 AGGGGAGAGATTGGGAAGACAATAGCAGGATCTGGGATCGGGATCGGCTTATGGAAC 4068
Qy 1061 ctgagggcgaaga----- 1074
Db 4069 CAGCTGGGCGCTCATCATGATGTATGACTGGCGCGCGCATCCGTCGAGACTTGGCGTAA 4128
Qy 1075 ----- 1074
Db 4129 TCATGTCATAGCTGTTCTGTTGTAATTTGTTATCCGCTCACAATTTCCACACACATA 4188
Qy 1075 ----- 1074

```



```

Oy 1929 gggcttgaacctgagtgaaacgaactaacgttaaggatttgcatactagcggatac 1988
      |||||
Db 5149 gggcttgaacctgagtgaaacgaactaacgttgaaggatttgcatactagatca 5208
      |||||

Oy 1989 atattgaaatgtaattaga 2007
      |||
Db 5209 AAAAGATCTTCACTAG 5227
      |||||

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 24, 2002, 01:01:13 ; Search time 5541.26 Seconds
(without alignments)
19765.864 Million cell updates/sec

Title: US-09-778-516A-1

Perfect score: 8115

Sequence: 1 gatgtacggcgccagatatac.....gcgtctgcctagcgacggtc 8115

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	775	9.6	954	9	AL044364 DKFP434C
2	772.6	9.5	1067	9	AU081137 AU081137
3	741.2	9.1	1089	9	AU081124 AU081124
4	719.2	8.9	724	9	AL645114 AL645114
5	708	8.7	800	9	AJ281449 AJ281449
6	703.6	8.7	715	9	AL661706 AL661706
7	702.2	8.7	707	9	AL656688 AL656688
8	700.6	8.6	1004	9	AJ281480 AJ281480
9	698.6	8.6	705	9	AL635845 AL635845
10	698.6	8.6	705	9	AL643164 AL643164
11	694.8	8.6	718	9	AL631067 AL631067
12	683.2	8.4	689	9	AL646532 AL646532
13	677.4	8.3	700	9	AJ281616 AJ281616
14	672.6	8.3	1070	9	AJ281552 AJ281552
15	669.2	8.2	675	9	AL636713 AL636713
16	665.2	8.2	670	9	AL635952 AL635952
17	663.2	8.2	669	9	AL660789 AL660789

18	663	8.2	703	9	AJ281437	AJ281437	4A3A-P4C3
19	662.6	8.2	672	9	AL640799	AL640799	
20	661.6	8.2	671	9	AL660035	AL660035	
21	659	8.1	670	9	AL642581	AL642581	
22	658.2	8.1	675	9	AL645322	AL645322	
23	652.2	8.0	657	9	AL641508	AL641508	
24	650.8	8.0	659	9	AL643220	AL643220	
25	649.8	8.0	680	9	AL646751	AL646751	
26	648	8.0	668	9	AL637190	AL637190	
27	646.4	8.0	808	9	AU176264	AU176264	
28	642.2	7.9	658	9	AL654693	AL654693	
29	642.2	7.9	659	9	AL647428	AL647428	
30	641.2	7.9	658	9	AL656146	AL656146	
31	636.6	7.8	651	9	AL660958	AL660958	
32	633	7.8	671	9	AL649526	AL649526	
33	632.4	7.8	718	12	AG105205	AG105205	
34	632	7.8	719	12	AG089715	AG089715	
35	628.6	7.7	646	9	AL633911	AL633911	
36	621.8	7.7	670	9	AL662130	AL662130	
37	621.4	7.7	734	9	AL039459	AL039459	
38	621.2	7.7	650	9	AL639359	AL639359	
39	615.2	7.6	718	12	AG010489	AG010489	
40	611.2	7.5	638	9	AL642392	AL642392	
41	610	7.5	695	12	AG100045	AG100045	
42	608.8	7.5	632	9	AL638320	AL638320	
43	606.8	7.5	680	12	BH235176	BH235176	
44	606.8	7.5	687	9	AU001481	AU001481	
45	606.2	7.5	664	9	AL650636	AL650636	

ALIGNMENTS

RESULT 1
LOCUS AL044364 954 bp mRNA linear EST 29-FEB-2000
DEFINITION DKFP434C172_s1 434 (synonym: htes3) Homo sapiens CDNA clone
ACCESSION DKFP434C172 3', mRNA sequence.
VERSION AL044364
KEYWORDS AL044364.1 GI:5432586
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 954)
Ansoerge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
EST (Ansoerge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Ansoerge W

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.
No r1 sequence available.
This clone (DKFP434C172) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubneweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de.

FEATURES

source
1..954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFP434C172"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

BASE COUNT	214 a	273 c	253 g	214 t
ORIGIN				
Query Match	9.6%;	Score 775;	DB 9;	Length 954;
Best Local Similarity	96.4%;	Pred. No. 2.8e-182;		
Matches 793;	Conservative	0;	Mismatches 30;	Indels 0;
Gaps				0;
OY	1031 atgtctgggagatgctgttggctctatgctctctgagtcggaagaaccagctgcatatg	1090		
DB	132 AAGGAGTCAACCTAAATGCGCTTCGCCCTCCGGCGCAATTGCAAGCTCTCATTTAAAG	191		
OY	1091 aatcgccaacgacgagggagagagcggtttgattatgagtcgctcttcgcgttcctgct	1150		
DB	192 AATCGGCCAACCGCGGGGAGAGGGGTTTGGCTATTGGGCCCTCTTCGGCTTCCCTG	251		
OY	1151 cactgactgcctgctgcgtcgtctgctgcgtgcgagcgatcaagctcaaaagc	1210		
DB	252 CACTACTTCGCGCCCTCGCTGCTTCGGCTCGCGGAGCGGTATCAGCTCATCTCAAGGC	311		
OY	1211 ggtatagcgttatccaaagaatcaaggagatagcgaagaagaacatgttgcaaaag	1270		
DB	312 GGTAAATACGTTATCCACAGAAATCAGGGGATATACGAGAAAGAACATGTGACAAAG	371		
OY	1271 cgaagcaaaagcgcaagaaaccgcaaaaagcgcgctgtgcggtttttccaaagctcgc	1330		
DB	372 CCAGCAAAAAGCGCAGAAACCGTAAAAAGGCCCGGTGCTGGCGTTTTCATATAGGCTCG	431		
OY	1331 ccccccctgacgagatcaacaaaatccgaagctcaagtcagatggtgycgaaccgcagag	1390		
DB	432 CCCCCCTGACGAGCATACAAAAAATTCAGACGCTCAATGATCAGAGTGGCGAAACCCACAG	491		
OY	1391 actataaagatccagagcgcttcccccctggaagctcccccctgcgcctccctgtccag	1450		
DB	492 ACTATTAAGATATCACAGGGGCTTCCCCCTGGAAGCTCCCGAGGCTCTCCGTTCGAC	551		
OY	1451 cctgcgcgttaacggatatacctgtccgcctttcccttcgggaagcgtaggcctttcca	1510		
DB	552 CCGTGGCGCTTACGGATATCTGCGCCCTTTCCTTCGGGAACGTTGGCCTTTCTCA	611		
OY	1511 atgcccacgctgtaagtatcagttcagttcggtgtgtaagtcgttcctcaagctggagctg	1570		
DB	612 TAGCTCACCGCTTAGGTATCTCAGTTCCGTTAGTGTGCTTCCTCCMAAGCTGGGCTGT	671		
OY	1571 gcaagaaccccccgcttcagcccgacgcgctgcgcttalcgcgtlaactacgtcttgatc	1630		
DB	672 GCACGAACCCCCCGCTTCAGCCCGACCGCGCCCTTATCCGGTATCATGCTTGAGAC	731		
OY	1631 caaccgcgttaagaacgactatgcgcactgacgaagcactgttaacagattagcag	1690		
DB	732 CAACCCGGTAGACACAGACTTATCGCCACTGGCAGCAGCACTGGTAACAGATTAAGCAG	791		
OY	1691 agcgaagatagtagcgctgctacagagttcttgaagtggtgcttaactacagctaac	1750		
DB	792 AGCGAGGTATGTAGCGCGGTCTACAGAGTCTTGAAGTGTGGCTTAACTAGCGCTACAC	851		
OY	1751 tagaagcgacgaatttgttatctgcgtctgctgttgtaagcagatlaactcttgaaaaagat	1810		
DB	852 TAGAAGAACAGATTATTTGGATCTGGCTGTGTGAAGCCAGTTACCTTGGAAGAAAGAT	911		
OY	1811 tgttagctcttgatccggaacaaacacacgcgctgttagcgt	1853		
DB	912 TGGTAGCTCTTGATCCGGCAAAACACACCGCTGTATCCGT	954		
RESULT 2				
LOCUS	AU081137	1067 bp	mRNA	linear
DEFINITION	AU081137 Oncohrynchus mykiss kidney infected by infectious hematopoietic necrosis virus Oncohrynchus mykiss cDNA clone K12, mRNA sequence.			
ACCESSION	AU081137			
VERSION	AU081137.1	GI:6431485		

KEYWORDS	EST.
SOURCE	Rainbow trout.
ORGANISM	Oncorhynchus mykiss
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Telostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
AUTHORS	Kono,T., Sakai,M. and LaPatra,S.E.
TITLE	1 (bases 1 to 1067)
JOURNAL	Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncorhynchus mykiss) Infected with Infectious Hematopoietic Necrosis Virus
COMMENT	Mar. Biotechnol. 2 (5), 493-498 (2001) Contact: Masahiro Sakai Faculty of Agriculture Miyazaki University 1-1 nishi yakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan Email: m.sakai@cc.miyazaki-u.ac.jp.
FEATURES	Location/Qualifiers
SOURCE	1..1067 /organism="Oncorhynchus mykiss" /db_xref="taxon:8022" /clone="K12" /clone_11b="Oncorhynchus mykiss Kidney Infected by Infectious Hematopoietic Necrosis Virus" /tissue_type="Kidney Infected by Infectious Hematopoietic Necrosis Virus"
BASE COUNT	235 a 292 c 288 g 250 t 2 others
ORIGIN	
Query Match	9.5%, Score 772.6, DB 9, Length 1067;
Best Local Similarity	98.6%; Pred. No.1.2e-181;
Matches 800; Conservative	0; Mismatches 9; Indels 2; Gaps 2;
QY 1076	ccagctgcattatataatcggcaacacgcgagggagagcggttcglatatggcgctc 1135
Db 259	ccagctgcattatataatcggcaacacgcgagggagagcggttcglatatggcgctc 318
QY 1136	ttccgccttcctgcgcacatgactgcgtgcgtcgtcgttcgtctgcgagcgatc 1195
Db 319	ttccgccttcctgcgcacatgactgcgtgcgtcgtcgttcgtctgcgagcgatc 378
QY 1196	agctcaactcaaaagcggtatatacggttatccacgaatcagggagatcagcagaa 1255
Db 379	agctcaactcaaaagcggtatatacggttatccacgaatcagggagatcagcagaa 438
QY 1256	catgtgagcaaaagcgcaaaagcgcaagacgtataaaagcgcgctgtgtgcgt 1315
Db 439	catgtgagcaaaagcgcaaaagcgcaagacgtataaaagcgcgctgtgtgcgt 498
QY 1316	tttccataagctcgcgccttcgacgagacatcaaaaatcgagcgtcaagtcaaggt 1375
Db 499	tttccataagctcgcgccttcgacgagacatcaaaaatcgacgctcaagtcaaggt 558
QY 1376	gcgaaccccgcaaggaactataagatacagcggtttccccccttgaaagctccctgcgt 1435
Db 559	gcgaaccccgcaaggaactataagatacagcggtttccccccttgaaagctccctgcgt 618
QY 1436	cctccctgttccgaacctgcgcttaccggatacctgttcgcgtcttcccttcgggaag 1495
Db 619	cctccctgttccgaacctgcgcttaccggatacctgttcgcgtcttcccttcgggaag 678
QY 1496	cgtgacgcttttcaatgctcaagctgtatagatataagttcgtgtgtagtggtgcgtc 1555
Db 679	cgtgacgcttttcaatgctcaagctgtatagatataagttcgtgtgtagtggtgcgtc 738
QY 1556	caagctggagctgtggaacgaacccccggttaagccgcaagcgcgtcgcgttacgcgttaa 1615
Db 739	caagctggagctgtggaacgaacccccggttaagccgcaagcgcgtcgcgttacgcgttaa 798
QY 1616	ctatcgcttcttgatccaaccggttaagacgaactatctgcacatgtgcagcaagcactgg 1675
Db 799	ctatcgcttcttgatccaaccggttaagacgaactatctgcacatgtgcagcaagcactgg 858

OY	1676	taacaggtttaacgaagcgaggatgatgtagggcgtgcctacaaggtcttctaagtgtgacc	1735
Db	859	TAAcAGGATTAACAAGAAGCGAGGTATGTAGGGCCGTCTCAAGATTCTTAAATGGTGCGC	918
OY	1736	tactactggtctaacataagaagacaglatlcttgatcttcgcgtctgcttgaagccagttac	1795
Db	919	TAACTACGGCTCACTAGTAGAAGACAGTATTGTGTACTTCGCCGTGCCTGAACCGACGTTAC	978
OY	1796	cttcggaaaaaagatgttgttagctcttgcattgcggcaaaacaaccacgcgtgtgtagcgttg	1855
Db	979	CTTTCGAAAAAAGATTGGTAGCTCTTGATCCGGCAAAACAA-CAAACGCTGTAGCGTGT	1037
OY	1856	ttttctgttcgaagcagcaattacgcgc	1886
Db	1038	TTTTTTTGTTC-AGCACACAGATTACC GCC	1067
RESULT	3		
AU081124			
LOCUS		1089 bp	mRNA linear EST 30-AUG-2001
DEFINITION	AU081124 Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus	Oncorhynchus mykiss cDNA clone KG'12,	
ACCESSION	AU081124		
VERSION	AU081124.1	GI:6431472	
KEYWORDS	EST.		
SOURCE	rainbow trout.		
ORGANISM	Oncorhynchus mykiss		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.		
AUTHORS	Kono,T., Sakai,M. and Lapetra,S.E.		
TITLE	Expressed Sequence Tag Analysis of Kidney and Gill Tissues from Rainbow Trout (Oncorhynchus Mykiss) Infected With Infectious Hematopoietic Necrosis Virus		
JOURNAL	Mar. Biotechnol. 2 (5), 493-498 (2001)		
COMMENT	Contact: Masahiro Sakai Faculty of Agriculture Miyazaki University 1-1 nishi gakuenkibanadai, Miyazaki, Miyazaki 889-2192, Japan Email: m.sakai@cc.miyazaki-u.ac.jp.		
FEATURES			
source	location/Qualifiers		
	1..1089		
	/organism="Oncorhynchus mykiss"		
	/db_xref="taxon:8022"		
	/clone="KG'12"		
	/clone_11b-"Oncorhynchus mykiss Kidney infected by infectious hematopoietic necrosis virus"		
	/tissue-type="Kidney infected by infectious hematopoietic necrosis virus"		
BASE COUNT	245 a 297 c 295 g 251 t	1 others	
ORIGIN			
Query Match	9.1%; Score 741.2; DB 9; Length 1089;		
Best Local Similarity	96.5%; Pred. No. 8.3e-174;		
Matches	801; Conservative 0; Mismatches 23; Indels 6; Gaps 4;		
OY	1076	ccaagtcattaatgaatcgcgccaacgcgcggggaagcgcgttgcglatatggcgcctc	1135
Db	260	CCAGCTGCATTAAATGAATCGCCAACGCGGGGAGAGCGGTTCGATTTGGCGCGTC	319
OY	1136	tttcgcttcctcgcgccatctgctgcgtcgcgttcgcttgcgcgcgcgcgcgcgcgttatc	1195
Db	320	TTTCGCTTCCTCGCTACCTGACTCCCTCGCTCGGTCTGTTGGCTGCGCGAGCGGTATC	379
OY	1196	agctcaaccaaaagcgcgtaatacgtttlccacaagaatcacgggagataacgcagsgaaa	1255
Db	380	AGCTCATCAACAAAGCGGTATACGTTATTCACACAAATCAGSGGATTAACGACGAAAAGA	439
OY	1256	cattgtgagcaaaagccagcaaaagccaggaacccgtlaaaaaagccgcgttgtgcgctt	1315

Db	Accession	Version	Source	Organism	Reference	Title	Journal	Comment
Db	440	CATGACGACAAAGGCCAGCAAAAGGCCAGAAACCGTAAAGGCCGCTTCTGGCCTT	499					
QY	1316	ttccatagatgcctcgcccccttgacgagcatcacaanaatcgacgctcaagtcaagtg	1375					
Db	500	TTTTCATGAGGCTCCGCCCCCTCGACGAGCATCAAAAATCAAGCTCAAGTCAGAGTG	559					
QY	1376	gcgaaacccgcagagactataaagttacagagcgtttcccccttgaaagtcctcgttg	1435					
Db	560	GGGAAACCCGACAGGACTATATAAGATACCAAGGCGTTTCCCTCGAAGCTCCCTCTGGC	619					
QY	1436	ctctcccttctccagaccctgcgctcttaccggaatacctgtctccgcttctccttcggaag	1495					
Db	620	CTCTCTCTTCTTCGACACCTCGCCCTTACCGGATACCTGTCCGCTTCTCTCCCTTCGGGAAG	679					
QY	1496	cgtgagccttctcgaatgctcaagctgtagtattcgaattcgaattgtagtgcgttc	1555					
Db	680	CGTGACGCTTCTCTCTACTACTCAGCTGTAGTATCTCACTTGGTGTAGTGTGCTGCTC	739					
QY	1556	caagcttgagcgtgtgagcagcaaccccccgcttaagcccgacgctgcgcttctcgttata	1615					
Db	740	CAAGTGTGGGCTGTGTGACAGAAACCCCGCTTACGCCCGACCGCTGTGCGCTTATCCGGTTAA	799					
QY	1616	ctatcgtctttagtcaacaccggttaagacaagactattatgcacatggcagcagcagctg	1675					
Db	800	CTATGCTCTTGAAGTCCAAACCCGGTAAAGACAGACACTTATGCCACTGTGACAGACCACTGG	859					
QY	1676	taacaggaattagcagagcgaagtatgtagcggtgctacagagttcttgaag-tgtgtgc	1734					
Db	860	TAAACGAGATTTAGCAAGAGCGAGGTATGTAGCGCGTCTACAGAGTTCTTGAATTTGTGTGC	919					
QY	1735	ctaactaagcgtctacactagaaggaagaatatttggatattcgtgccttgcctggaagccagtta	1794					
Db	920	CTAATCTAGCGCTACCTCTGAAGAAAGATATTTTGTATCTGCCCTCTGTGAGCCAGTTTA	979					
QY	1795	ccttcggaagaaagattgtgtagctctt-gatccggaacaacaaca---cgctgtgtagc	1850					
Db	980	CCTTCGGAAAAAGAGTGGGTAGCTTTTGATTCGGGCAAAACCAACACCGCTGGGAGAGG	1039					
QY	1851	ggtgtttttttgttttgccaagca-gcaattacgcgcagaaaaaagaat	1899					
Db	1040	GGGTCTCTTTTGTGTTCAGACGACGACAGATTTACCCGCGGAAAAAAGCGAT	1089					
RESULT	4							
AL645114		724 bp	mRNA	linear	EST	12-DEC-2001		
LOCUS	AL645114							
DEFINITION	AL645114 XGC-neurula Silurana tropicalis cDNA clone TNeu017b19 5',							
ACCESSION	AL645114							
VERSION	AL645114.1							
KEYWORDS	EST.							
SOURCE	western clawed frog.							
ORGANISM	Silurana tropicalis							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
AUTHORS	Amphibia; Batrachia; Anura; Mesodactylia; Pipidae;							
TITLE	Xenopodinae; Silurana.							
JOURNAL	1 (bases 1 to 724)							
COMMENT	Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.							
	Sanger Xenopus tropicalis EST project 2001 (10_2001)							
	Unpublished (2001)							
	Contact: Huckle E							
	Sanger Centre							
	Hinxton, Cambridgeshire, CB10 1SA, UK							
	Email: tropesanger.ac.uk							
	Sanger Xenopus tropicalis EST project 2001							
	TROPICALIS_SEQUENCE_ID: TNeu017b19.sp6							
	Sequencing primer: SP6							
	This sequence is from a Xenopus Gene Collection (XGC) library							
	constructed by Aaron M. Zorn.							
	Location/Qualifiers							
	1..724							
FEATURES								
source	/organism="Silurana tropicalis"							


```

|||||
Db 372 CGCTTGAAGTCCACCCGCGTAAGACACGACTTATCGCCACTG6CACACCACTGGTAAC 431
|||
Qy 1680 agagattagagagcagagatgtaggcggtgctacaagaattcttgaagcgtgagcctaac 1739
|||||
Db 432 AGGATTAGCAGAGCAGAGGTATGAGCGGCTGCTACAGAGTCTTGAGAGTGCGGCTTAAC 491
|||
Qy 1740 tacgcgtacacagagagcagatttggtatctgcctgcctcgtgaagcagttacttc 1799
|||||
Db 492 TACGGCTACACTAGAAAGAACAGTATTGGTATCTGCGCTGCTGAGACCACTTACCTTC 551
|||||
Qy 1800 ggaataaagagttgtagctcttgatccgcgcgaacaacacacgcgtgtagcgtgtttt 1859
552 GGAATAAAGAGTTGTAAGTCTTGATCCGCGAACAACACCGCTGTAGCGGCTGTTT 611
|||
Qy 1860 ttgtgttgaaagcagcagattacgcgcgaataaagagttcagaagatccttgatc 1919
|||||
Db 612 TTTGTTTGCAGACACAGATTACGGCGAGAAAAAAGAGATCTCAAGAGATCCTTTGATC 671
|||
Qy 1920 ttctacagggctcgaagcctcagtggaacgaatacctcgttaagagatttgatcgt 1979
|||||
Db 672 TTTTCTACGGGCTCTGACCTCAGTGGAGAACGAAACTCAGTTAAGGATTTTGGTCATG 731
|||
Qy 1980 agcggatacatatttgaatgtaattga 2007
|||
Db 732 AGATTATCAAAAAAGATCTTACACTAGA 759

```

RESULT 6

```

AL661706 715 bp mRNA linear EST 13-DEC-2001
LOCUS AL661706 XGC-neurula Silurana tropicalis cDNA clone TNeu049m20 5',
DEFINITION mRNA sequence.
ACCESSION AL661706
VERSION AL661706.1 GI:17676434
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 715)
AUTHORS Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu049m20.sp6
Sequencing primer: Sp6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1..715
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu049m20"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCS107; Site 1: EcoRI; Site 2: NotI; cDNA
was oligo dt primed from 5' end of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 154 a 209 c 196 g 156 t
ORIGIN

```

```

Query Match 8.7%; Score 703.6; DB 9; Length 715;
Best Local Similarity 99.4%; Pred. No. 1.8e-164;
Matches 706; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```

```

Qy 1076 ccagctcatlaatgatcggccaacgacgagcgtgttcgtatctgagcgtc 1135
|||||
Db 6 CCAGCTGATTAATGATCGGCCAACGCCGGGGAAGAGGGTTTGCGATTGGCGCTC 65
|||||
Qy 1136 ttccgcttcctcgtcactgaactgctgcgtgcgttcgttcgttcgttcgttcgttcgttc 1195
66 TTCCGCTTCCTCGCTCAGTGAAGTCCGCTGCGCTGCGTTCGCTGCGGCGAGCGGATTC 125
|||||
Qy 1196 agctcaactcaaaagcgttaatacgtttacaaagaatacagggataaagcaggaagaa 1255
126 AGCTCACTCAAAAGCCGTATATACGTTATCCACAGATCAGGGGATACGCAAGAAAGAA 185
|||||
Qy 1256 catgtgagcaaaaagcgaacaaagccaggaacgtaaaaagccgctgttcgttcgttc 1315
186 CATGTAGCAAAAAGCCGACGAAAGCCGAGAACCTTAAAAAGCCGCTGTGCGCTT 245
|||||
Qy 1316 ttccataagctcgcgcgccttcgacgagatcacaaaatcgaagctcaagtcagagtg 1375
246 TTTCCATAGGCTCCGCCCTCAGCAGCATCAAAAAATCGACGCTCAAGTCAGAGTG 305
|||||
Qy 1376 ggcgaacccgacagactataaagatacagcgtttccctcgtgaagctcctcgttcg 1435
306 GCGAAACCCGACAGGACTATAAAGATACAGGCGTTTCCCTCGAAGCTCCCTCGTGC 365
|||||
Qy 1436 ctctcctcttcgagccctcgcgttcacgagatccctgcgttcctccttccttcgggag 1495
366 CTCTCCTGTTCCGACCTGCGCTTACCGGATACGTGTCGCTTCTCCCTTCGGGAG 425
|||||
Qy 1496 cgtgcgcttcctcaatgctcaacgctgtagtaccagttcgttgtagtgcgttc 1555
426 CGTGGCGCTTTCATAGCTCAGCTGTAGTATCTCAGTGGGTAGTGGTTCGCTTC 485
|||||
Qy 1556 caagctggcgtgtgtgcagcaaccccggttcagccgacgctgcgttcctacgtgta 1615
486 CAAGCTGGGCTGTGTGACGAAACCCCGCTTACGCCGCTGCGCTTATCCGTTAA 545
|||||
Qy 1616 ctatgcttcgttgatccaacccggttaagacagccttatccgcttcgacgacactgt 1675
546 CTATGCTGTTGATCCACCCGGTAAACACGACTTATGCGCACTGGACACCCACTGG 605
|||||
Qy 1676 taacagattagcagagcagagatgtagcgtgtgtacagagttcttgaagtgtgtgc 1735
606 TAACAGATTAGCAGAGCGAGGTATGAGCGGCTGTACAGATTCTTGAAGTGTGTGGCC 665
|||||
Qy 1736 taactacggtcactaagaagcagatttggtatctgcctcgtcgtca 1785
666 TAACCTACCGCTACACTAGAAAGACAGTATTGGTATCTGCGCTCTGCTGA 715
|||||
Db

```

RESULT 7

```

AL656688 707 bp mRNA linear EST 13-DEC-2001
LOCUS AL656688 XGC-neurula Silurana tropicalis cDNA clone TNeu032h21 5',
DEFINITION mRNA sequence.
ACCESSION AL656688
VERSION AL656688.1 GI:17669128
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 707)
AUTHORS Huckle,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J.
TITLE Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu032h21.sp6

```

```

BASE COUNT 154 a 209 c 196 g 156 t
ORIGIN

```

Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
Location/Qualifiers
1. 707
source

/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu032h21"
/clone_lib="Xgc-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: pCS107; Site.1: EcoRI; Site.2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into pCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 153 a 205 c 195 g 154 t
ORIGIN

Query Match 8.7%; Score 702.2; DB 9; Length 707;
Best Local Similarity 99.6%; Pred. No. 3.9e-164;
Matches 704; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1079 gacgcaataatgaatgcagcccaacgcgcgggagagagcggttcgattggcgctcttc 1138
|||||
Db 1 GCTGCAATTATGATGCGGCAACGCCGGGAGAGCGGTTTCCGTAATGGCGCTCTTC 60
Qy 1139 cgcttcctcgctactgactgctgctgctgctgctgctgctgctgctgctgctgct 1198
|||||
Db 61 CGCTTCTCGCTCAGTACTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Qy 1199 tcaactaaagcggtataatcggtatccaaagaatacgaagagataagcagaagaacat 1258
|||||
Db 121 TCAGCTAAAGGCGGTATAGGTTATCCACAGATACAGGGGATTAACAGAAAGAACAT 180
Qy 1259 gtagcaaaagcgcaagaaagcgcaagaaacgtaaaagcgcgctgctgctgctgctt 1318
|||||
Db 181 GTGAGCAAAAGGCGCACCAAAAGGCCAGAACCTTAATAAGCCGCTGCTGCTGCTTTT 240
Qy 1319 ccaatagctcgcgcgcctcgaagcatcaaaaatcgacgctcaagtcagaagtgctg 1378
|||||
Db 241 CCATAGGCTCCGCCGCCCTGACGACATCAAAAATTCAGCGCTCAAGTCAGAGGTGGC 300
Qy 1379 aaaccgcaagagctataaagataacagagcggttccccctggaaagctcctgctg 1438
|||||
Db 301 AAACCCGACAGACTATAAAGATACAGAGCGTTTCCCTCGAAGCTCCCTCGCGCTC 360
Qy 1439 tccgttcgcaacctcgctgctacagataccgttcgcaccttcctccttcggaagcgt 1498
|||||
Db 361 TCCTGTTCCGACCTCGCTTACCGGATTCCTGCGCTTCTCCCTTCGGGAAAGCT 420
Qy 1499 ggcgcttctcaatgctcaagctgtatgtaatcagttcgggtgtaagtcgttcgctca 1558
|||||
Db 421 GGGGCTTTTCATAGCTCAAGCTGTAGTATCTCAGTTCCGTTAGAGTCTGTCGCTCAA 480
Qy 1559 gctggagctgtgtgcaagaacccccgcttcagccgcaagcgtcgcttccttcggtta 1618
|||||
Db 481 GCTGGGCTGTGTGTCAGAACCCCGCTTCAGCCCGAAGCGCTGCGCTTATCCGGTAAC 540
Qy 1619 tgccttcgagtcgaaccggttaagacagactatcgccactggcagcagcagctgtaa 1678
|||||
Db 541 TCGTCTTGAGTCCAAACCCGGTAAAGACAGACTTATCCACTGACAGACCACTGGTAA 600
Qy 1679 caggaatgaagagcaggtatgtaagcagtgctacagagttcctgaagtcgtggcctaa 1738
|||||
Db 601 CAGGATTACAGAGCAGGATGTAGGCGGTGCTACAGATTTCTTAAGAGTGCGCTTAA 660
Qy 1739 ctaaggtacactagaagaagcagtaattggtatcgtcgctcgtcgtaa 1785
|||||
Db 661 CTACGGCTACACTAGAAAGACAGTATTGTTGATCTGCGCTGCTGA 707

RESULT 8

AJ281480
LOCUS 1004 bp mRNA linear EST 30-JUN-2000
DEFINITION 4A3A-P4G8-F Anopheles gambiae immune competent 4A3A Anopheles
gambiae cDNA clone 4A3A-P4G8, mRNA sequence.
ACCESSION AJ281480
VERSION AJ281480.1 GI:6929360
KEYWORDS EST.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae
; Anophelinae
1 (bases 1 to 1004)
DIMPOULOS, G., CASAVANT, T.L., CHANG, S., SCHELTZ, T., ROBERTS, C.,
DONOHUE, M., SCHULTZ, D., BENES, V., BORK, P., ANSORGE, M., SOARES, M. B.
and KAFATOS, F. C.
Anopheles gambiae pilot gene discovery project: identification of
mosquito innate immunity genes from expressed sequence tags
generated from immune-competent cell lines
Proc. Natl. Acad. Sci. U.S.A. 97 (12), 6619-6624 (2000)
20300950
Contact: Dimopoulos G
Folis C. Kafatos Laboratory
European Molecular Biology Laboratory
Meyerhofstrasse 1, 69117 Heidelberg, Germany.
Location/Qualifiers
1. 1004
/organism="Anopheles gambiae"
/strain="4A r/x"
/db_xref="taxon:7165"
/clone="4A3A-P4G8"
/clone_lib="Anopheles gambiae immune competent 4A3A"
/cell_line="Immune competent 4A3A"
/lab_host="E. coli DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site.1: EcoRI; Site.2: NotI; sequenced from
forward priming site which reads from the 3' end of the
cDNA. The 4A3A is a directionally cloned and normalized
cDNA library that was constructed from the 4A3A cell line
oligo-T primed cDNA according to: Bonaldo, Lennon & Soares
(1996) : Normalization and Subtraction: Two approaches to
Facilitate Gene Discovery, Genome Research 6, 791-806."

BASE COUNT 252 a 262 c 244 g 244 t 2 others
ORIGIN

Query Match 8.6%; Score 700.6; DB 9; Length 1004;
Best Local Similarity 97.3%; Pred. No. 1.1e-163;
Matches 712; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1276 aaaaagccaggaacgtaaaagcgcggttcgtgcttccatagagctccgccccc 1335
|||||
Db 1 AAAAGGCGACAGAACCGTAATAAGGCGCGTGTGCGCTTTTCCATAGAGCTCGGCCCC 60
Qy 1336 ctgacgagcatcaaaaatcgacgctcaagtaaggtgagaaaccgacagagactat 1395
|||||
Db 61 CTGACGAGCATCAAAAATGACGCTCAAGTAGAGGTGGCAACCCGACAGAGATAT 120
Qy 1396 aaagatcacagcggttccccctggaaagctccctcgcgctcctcgttcgcaacctgc 1455
|||||
Db 121 AAAGATACCAAGGCGTTTCCCGCTGGAAGCTCCTCGGCGCTCTCTGTTCCGACCTGC 180
Qy 1456 cgcttacgatacctgctcgaccttcccttccttcggaagcgtggcgcttccaatgct 1515
|||||
Db 181 CGCTTACCGGATACCTGTCGCGCTTTCCTTCGGAACGAGGCGCTTTCATATAGCT 240
Qy 1516 caagctgtaagtatcagttcgttgtaagtgcttcgaagcgtggcgctgtgtgcaag 1575
|||||
Db 241 CACGCTGTAGATATCTCAGTTCCGTAGGTGCTGCTCCCTCAAGCTGCGCTGTTGACAG 300
Qy 1576 aaccgccggttcagccgcaagcgttcgcttaccggttaactatcgcttgagtcgaac 1635
|||||
Db 301 AACCCCGCTTACAGCCCGACCGCTGCGCTTATCCGTAATCTGCTTGAATCCAAC 360

QY	1636	cggtgaagcaacgactatctgcacacgctgagcagcagccacgtgtaacaggatttagcaaaagca	1695
Db	361	CGGTAAAGACGACATTATCCGCTACGACGGACCCACGTGTAAAGAGATTACCAAGCCA	420
QY	1696	ggtatctagcgctgtctacagaagttctcttgaagtggtgtgacctactacgctcactagaa	1755
Db	421	GGTATGTAGCGCGCTCCCTACAGAGTTCTTGAAATGGTGGCCTTACTACGGCTTACACTAGAA	480
QY	1756	ggaacgattcttgtaactctgcgcctctgctgtaagccagttactcttcgnaaaagattgta	1815
Db	481	GAACAGATTGTTGGTATCTGCGCTGTGTAAGCCAGTTACTCTCGAAAAAGATTGGTA	540
QY	1816	gctcttgatccgycgaacaaacacacccgctgtgtagcgtgtgttcttctgttgcgaagcagc	1875
Db	541	GCTCTTGATCCGGCCAAACCAACCCGCTGGTGAAGGGTGCTTTTGTTCGAAAGCAGC	600
QY	1876	agattacgcgcgaaacaaagatcctcaagaagatcccttgaactcttctctaaggggtctg	1935
Db	601	AGATTACCCGCAGAAAAAAGATCTCAGAAGATCCCTTGATCTTTTCTACGGGGTCTG	660
QY	1936	acgcctcagtggaacgaaacactcagttaaggatttgcctactgaacgatacatatlyg	1995
Db	661	ACGCCTACGTGAGACAAATCATCCTTAAGGATTTTGTCATAGAGATTATCAAAAAGAA	720
QY	1996	aatgtattataga 2007	
Db	721	TCTTCACCTTAA 732	

RESULT	9
AL635845	
LOCUS	705 bp mRNA
DEFINITION	linear EST 12-DEC-2001
ACCESSION	AF635845 XGC-neurula silurana tropicalis cDNA clone tne0015811 5'
VERSION	AL635845 mRNA sequence.
KEYWORDS	AL635845.1 GI:16787824
SOURCE	EST.
ORGANISM	western clawed frog. silurana tropicalis

REFERENCE	(bases 1 to 705)
AUTHORS	Huckle, E., Taylor, R., Ashurst, J. L., Zorn, A. M. and Rogers, J
TITLE	Samner Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL	Unpublished (2001)
COMMENT	Contact: Huckle E

This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
location/Qualifiers
1..705
source

```

/lab_host="Escherichia coli DH10B"
/notes=Vector: PCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from neurina
EcoRI-NotI cut cDNA was then ligated into PCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT      152 a      205 c      193 g      155 t
ORIGIN

```

Query Match	8.68; Score 698.6; DB 9; Length 705;
-------------	--------------------------------------

Best Local Similarity 99.4%; Pred. No. 3,1e-163;
Matches 701; Conservative 0; Mismatches 4; Indels 0; Gaps 0

[illegible]

RESULT	10
AL643164	
LOCUS	
DEFINITION	AT643164 705 bp mRNA linear EST_12-DEC-2001 AL643164 XGC-neurula Silurana tropicalis cDNA clone TNeu017ml3 5', mRNA sequence.
ACCESSION	AL643164
VERSION	AL643164.1 GI:16795289
KEYWORDS	EST.
SOURCE	western clawed frog. Silurana tropicalis
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Silurana. 1 (bases 1 to 705)
REFERENCE	Huckie,E., Taylor,R., Ashurst,J.L., Zorn,A.M. and Rogers,J. AUTHORS Sanger xenopus tropicalis EST project 2001 (10_2001) TITLE unpublished (2001) JOURNAL Contact: Huckie E COMMENT

Hinxton, Cambridgeshire, CB10 1SA, UK
Email: trop@sanger.ac.uk

Db 301 TTATCCGGTAACATATCTGTTGAGTCACAAACCCGGTAAGACACGACTTATGCCACTGGCA 360
Oy 1665 gcaagcactgtaaacagagattagcagagcgaagtatgtaggcggtgtctacagaattcttg 1724
Db 361 GCAGCCACTGTTACACAGGTTTAGCAGAGGAGGATGTAAGGGGGTCTACAGAGTTCTTG 420
Oy 1725 aagtggtgacctaacctacagctacactagaagaagacagatattgtgtatctgcgtctg 1784
Db 421 AAGTGTGGCTTAACCTAGGCGTACACTAGAACACAGATTTGGTATCTGCCCTCTGCTG 480
Oy 1785 aagcagcttacctcgcgaagaaagatgtgtagctcttgatccgcgaacaacaaccgct 1844
Db 481 AAGCCAGTACCTTCGGAAAAAGAGTTGTAGCTTGTATCCGGCAAAACCAACCCGCT 540
Oy 1845 gtagcagtggtttttgtttgttgcaagcagcagattacgcgcgaagaaaaaagatctaa 1904
Db 541 GGTACGGGTGGTTTTTGTGTGCAAGCAGCAGATTACCGCGAGAAAAAAGATCTCAA 600
Oy 1905 gaagatccttgatcttcttctacgcggtcgtcagctcagtggaacgaacacagcttaa 1964
Db 601 GAAGATCCTTGATCTTTTCTTACGGGGTCTGACGCTCAGTGAACGAAACTCAGCTTAA 660
Oy 1965 gggatttggtcagcagcgatatacttgatgattttaga 2007
Db 661 GGGATTTGTGTCATGATGATTAACAAGAGATCTTCACTTAGA 703

RESULT 15
AL636713 675 bp mRNA linear EST 12-DEC-2001
LOCUS AL636713 XGC-neurula Silurana tropicalis cDNA clone TNeu013j03 5',
DEFINITION mRNA sequence.
ACCESSION AL636713
VERSION AL636713.1 GI:16788692
KEYWORDS EST.
SOURCE western clawed frog.
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipridae;
Xenopodinae; Silurana.
1 (bases 1 to 675)
Huckle, E., Taylor, R., Ashurst, J.L., Zorn, A.M. and Rogers, J.
Sanger Xenopus tropicalis EST project 2001 (10_2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Huckle E
Sanger Centre
Hinxton, Cambridgeshire, CB10 1SA, UK
Email: tropesanger.ac.uk
Sanger Xenopus tropicalis EST project 2001
TROPICALIS_SEQUENCE_ID: TNeu013j03.sp6
Sequencing primer: SP6
This sequence is from a Xenopus Gene Collection (XGC) library
constructed by Aaron M. Zorn.
FEATURES
source Location/Qualifiers
1..675
/organism="Silurana tropicalis"
/db_xref="taxon:8364"
/clone="TNeu013j03"
/clone_lib="XGC-neurula"
/dev_stage="neurula"
/lab_host="Escherichia coli DH10B"
/note="Vector: PCS107; Site_1: EcoRI; Site_2: NotI; cDNA
was oligo dT primed from 5ug of poly A+ RNA from neurula.
EcoRI-NotI cut cDNA was then ligated into PCS107 with
EcoRI at the 5' end and NotI at the 3' end."
BASE COUNT 160 a 187 c 179 g 149 t
ORIGIN

Query Match 8.2% Score 669.2; DB 9; Length 675;
Best Local Similarity 99.6% Pred. No. 6.9e-156;
Matches 671; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1201 actcaaaagcgtaatacagttatccacagaatcaaggatatacgcaggaagaacatgt 1260
Db 2 ACTCAAAAGCGGTAAATACGTTATCCACAGATTCAGGGGATTAACGAGAAAGACATGT 61
Oy 1261 gagcaaaagccagcaaaagccaggaacccgtataaaagccggtgtgtcgtgtttcc 1320
Db 62 GAGCAAAAAGCCACAGCAAAAAGCCAGGAAACCGTAAAAAGCCCGCTGTGCGGTTTTC 121
Oy 1321 ataggtccgcccccttgacagatcacaataatcgacgtcgaagtcgaagtggtgcga 1380
Db 122 ATAGGCTCCGCCCTTGACAGACATCAAAAATCGAGCTCAAGTCAAGAGGTGGCGAA 181
Oy 1381 acccgaagaactataaagatacccaagggttcccccgtgaagctccctcgtgcgtc 1440
Db 182 ACCCGACAGACTATTAAGATACAGCGCTTCCCTTGAAAGCTCCCTGTCGCTC 241
Oy 1441 cgtttccgaacctgcgcgttaccgatacctgtccgcttctcccttcgggaagcgttg 1500
Db 242 CTGTTCCGACCTGCGCGTTACCGGATACCTGTCGCCCTTCTCCCTCGGGAAGGTGG 301
Oy 1501 cgtttctcaatgctcagcgtgtgaagttatctcagttcgtgtgaagtcgttcgaagc 1560
Db 302 CGCTTCTCATAGCTACACGCTGTAGGTCTAGTTCGAGTTCGAGTTCGCTCCAAAGC 361
Oy 1561 tggcgtgtgtgcagcaacccccgttcaagccgacccgtgcgccttatccggttaatac 1620
Db 362 TGGCGTGTGTCACGACACCCCGCTTACGCCGACCGCTGCGCTTATCCGTAATATC 421
Oy 1621 gttctgaagtcacaaccgcgtgaagacacgactatcgccactggaagcagcactgtgta 1680
Db 422 GTCTTGAGTCCAAACCCGTAAGACACAGACTTATCCGCACTGGCAGCAGCAGCTGTAA 481
Oy 1681 ggttagcaggaagcaggtatgtgaagcgtgtcgaagagttcttgaagtggtgctaa 1740
Db 482 GGATTAGCAGAGCGAGGTATGTAGCGGTGTACAGAGTTCTTGAAGTGTGGCCCTAACT 541
Oy 1741 acggtacactagaagaagacagatattgtgatactgcgctctgtgaagccagttactcg 1800
Db 542 ACGGCTACACTGAAAGAACAGATTTTGTGATCTGGCGCTGCTGAAGCCAGTTACTCTG 601
Oy 1801 gaaaaagagtggtagctcttgatccggcaaaacacacacgcgctgtgtagcgtgtttt 1860
Db 602 GAAAAAGAGTTGTAGCTCTTGTATCCGGCAAAACCAACCGCTGTGACGGTGTGTTT 661
Oy 1861 ttgtttgcaagcag 1874
Db 662 TTGTTTGCAGCAG 675

Search completed: May 24, 2002, 07:54:33
Job time: 24800 sec

